THESIS FOR THE DEGREE OF DOCTOR OF TECHNOLOGY

Robustness quantification in yeast

A methodology to study phenotypic, evolutionary, and genomic aspects of microbial robustness.

CECILIA TRIVELLIN

Department of Life Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2024

Robustness quantification in yeast
A methodology to study phenotypic, evolutionary, and genomic aspects of microbial robustness.
CECILIA TRIVELLIN
ISBN 978-91-7905-990-3

© CECILIA TRIVELLIN, 2024.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5456
ISSN 0346-718X

Division of Industrial Biotechnology
Department of Life Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover: The PhD journey through yeast performance and robustness data analysis
Artwork by: Cecilia Trivellin

Printed by Chalmers Digitaltryck,
Gothenburg, Sweden 2024

"Ricordate, tosa, che se te tegni la màn serà non te va fora gnente, ma no te entrerà mai gnente"

" *Remember, if your hand is closed, nothing can go out, but nothing can enter either*"

- Nonna

Preface

This dissertation partially fulfils the requirements for obtaining the degree of Doctor of Philosophy at the Department of Life Sciences at Chalmers University of Technology. The work was supported by the Novo Nordisk Foundation grant Distinguished Investigator 2019 - Research within biotechnology-based synthesis & production (grant # 0055044) awarded to Prof. Lisbeth Olsson. The PhD studies were carried out between November 2019 and April 2024 under the supervision of Prof. Lisbeth Olsson and the co-supervision of Peter Rugbjerg, PhD, with Prof. Carl Johan Franzén as examiner.

Most of the work in this thesis was carried out at the Division of Industrial Biotechnology at the Department of Life Sciences at Chalmers University of Technology. Evolution experiments and genome sequencing were conducted during a research visit at Harvard University in collaboration with Professor Michael Desai, supported by grants from the Barbro Osher Endowment and the Energy Area of Advance at Chalmers. Phenotypic assays were performed at Gothenburg University in collaboration with Dr. Karl Persson. Analysis of the genome sequences was performed in collaboration with Dr. Diana Ekman and Dr. Estelle Proux-Wéra from SciLifeLab, Stockholm funded by the long-term bioinformatic support granted by the Knut and Alice Wallenberg Foundation (KAW).

Cecilia Trivellin,
April 2024

Robustness quantification in yeast

A methodology to study phenotypic, evolutionary, and genomic aspects of microbial robustness

Cecilia Trivellin
Division of Industrial Biotechnology – Department of Life Sciences
Chalmers University of Technology

Abstract

Bioprocesses contributes to the shift towards a more sustainable economy. In bioprocesses, valuable chemicals can be generated from renewable resources while, at the same time, reducing carbon emissions. A major hurdle in bringing bio-based products to market is the time and cost involved in designing efficient cell factories. Cell factories developed in controlled laboratory settings achieve high yields and productivities, but often fail at a larger scale because of unforeseen perturbations. Microbial robustness, i.e., the ability to maintain functionality despite perturbations, is critical for designing cell factories but remains poorly studied, particularly with respect to quantification as well as evolutionary and genetic aspects.

In this thesis work, mathematical evaluation, phenotypic characterization, evolution and genomics were applied to address the lack of quantification methods and explore robustness in yeast. A Fano factor-based approach for measuring robustness across multiple parameters and perturbations was created. Measurement of physiological data revealed trade-offs between robustness and performance in yeast. Moreover, when screening yeast deletion libraries, it pointed to the MET28 gene, which encodes a transcription factor regulating sulfur metabolism, as a mediator of robustness. Finally, evolution in fluctuating environments improved robustness in the industrial strain Ethanol Red, but not in two laboratory strains, contrasting with fitness trends.

Altogether, applying robustness quantification to various experimental set-ups, enabled the identification of key genes and metabolic processes linked to enhanced robustness. This thesis thereby contributes to the field of physiology, particularly in the context of robustness. The developed techniques have potential to advance design optimization and testing of robust strains in laboratory settings, thereby enabling a faster scale-up to industrial environments.

Keywords: High-throughput, *Saccharomyces cerevisiae*, adaptive laboratory evolution, fluctuating conditions, bioproduction, perturbations

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I. Olsson*, L., Rugbjerg*, P., Torello Pianale*, L., & **Trivellin*, C.** (2022). Robustness: linking strain design to viable bioprocesses. *Trends in biotechnology*, *40*(8), 918–931. https://doi.org/10.1016/j.tibtech.2022.01.004

II. **Trivellin, C.**, Rugbjerg, P., & Olsson, L. (2022). Quantification of microbial robustness in yeast. *ACS Synth. Biol. 2022, 11, 4, 1686–1691*. https://doi.org/10.1021/acssynbio.1c00615

III. **Trivellin, C.**, Rugbjerg, P., & Olsson, L. (2023). Performance and robustness analysis reveals phenotypic trade-offs in yeast. Life Science Alliance. 2023, 7,1, e202302215. https://doi.org/10.26508/lsa.202302215

IV. **Trivellin, C.**, Torello Pianale, L., & Olsson, L. (2023). Robustness quantification of a mutant library screen revealed key genetic markers in yeast. (*Submitted to Microbial Cell Factories*)

V. **Trivellin, C.**, Ekman, D., Persson, K., Olsson L., & Desai, M. Evolution of microbial robustness in fluctuating environments. (*Manuscript*)

*The authors contributed equally

Contribution Summary

I. I contributed to group discussions that generated the key concepts of the study. All authors contributed to writing and editing of the manuscript.

II. I conceived the quantification formula together with PR and LO. I planned and performed the experimental work. I analyzed the data, interpreted the results, and wrote the manuscript. I edited the manuscript together with the other authors.

III. I planned and performed the experimental work. I analyzed the data, interpreted the results, and wrote the manuscript. I edited the manuscript together with the other authors.

IV. I planned the experiments, analyzed the data, carried out the experiments, and wrote the manuscript. LTP designed and constructed the deletion strains, as well as contributed to the interpretation of results. I edited the manuscript together with the other authors.

V. I conceived the idea, planned and carried out the evolution experiments, and prepared the library for sequencing. Along with KP, I performed the phenotypic assays. KP performed the pre-processing of phenotypic data. I analyzed the data on fitness and robustness. DE performed genome sequencing analysis; whereas I analyzed the variants data. Along with MD and LO, I interpreted the results. I wrote the manuscript, with input from DE for the genome sequence analysis procedure and KP for the phenotypic assays procedure. I edited the manuscript together with the other authors.

Table of Contents

Preface V

Abstract VII

List of Publications VIII

Contribution Summary IX

Table of Contents X

Abbreviations XII

Chapter 1. Introduction 1

1.1. Microbial robustness 1
1.2. Perturbations and robustness quantification 1
1.3. Intrinsic aspects of robustness 2
1.4. Aim of the thesis 2

Chapter 2. Microbial Performance and Perturbations in Bioprocesses 7

2.1 Sustainable production of commodities 7
2.1.1 The requirement for robust biocatalysts 9
2.2 Cell factories in bioprocesses 9
2.2.1 *Saccharomyces cerevisiae* 10
2.2.2 Performance metrics for the design of cell factories 12
2.3 Perturbations 17
2.3.1 Perturbations in lignocellulose fermentation processes 17
2.3.2 Bioprocess perturbations 21
2.3.3 Three aspects in perturbation studies important for performance assessment 22
2.4 Methods, datasets, and strains 27

Chapter 3. Microbial Robustness 33

3.1. The concept of robustness 33

3.1.1 Microbial robustness definition 33

3.1.2 Relevance of microbial robustness 34

3.2 Quantification of microbial robustness 35

3.2.1 State-of-the-art in robustness quantification 37

3.2.2 Fano factor-based quantification 37

3.2.3 Robustness and tolerance 40

3.2.4 Robustness quantification using available datasets 41

3.2.5 Combined effects of perturbations on robustness 43

3.3 Robustness vs performance 45

3.3.1 Trade-offs 47

3.3.2 Pareto fronts 49

Chapter 4. Microbial robustness concepts applied to evolution and genomics 53

4.1. Genetic markers of microbial robustness 53

4.1.1. Previously investigated markers 54

4.1.2 Yeast deletion collections: a tool for identifying robustness markers 54

4.2 Evolution of microbial robustness 57

4.2.1 Adaptive laboratory evolution 58

4.2.2 Phenotypic assays in evolution experiments 59

4.2.3 Fitness and robustness in ALE 60

4.2.4 Genomic variants 66

4.2.5 Robustness and history-dependent behavior 68

Chapter 5. Conclusions and Outlook 71

5.1 Conclusions 71

5.2 Outlook and Future perspectives 73

5.2.1 Microbial robustness quantification 73

5.2.2 Metabolic and genetic markers of robustness 75

Acknowledgements 77

References 81

Abbreviations

1G	First generation
2G	Second generation
AFCS	Automatic flight control system
ALE	Adaptive laboratory evolution
CDW	Cell dry weight
CV	Coefficient of variation
HMF	5-Hydroxymethylfurfural
LBCM	Laboratorio de biologia cellular e molecular
SDG	Sustainable Development Goal
TRY	Titer-rate-yield
YKO	Yeast knockout

Chapter 1. Introduction

Biological systems are constantly subjected to perturbations, which may arise internally, for example, via genetic mutations, or from external factors such as changes in temperature. The system's ability to withstand such perturbations and exhibit a stable phenotype is called microbial robustness (1,2). Robustness has been rarely quantified in applied research or when evaluating bioprocesses. The absence of practical quantification techniques hinders the exploration of the fundamental mechanisms that contribute to robustness.

The overall aim of my **thesis** was to develop a methodology for quantifying microbial robustness and to explore its use in different applications.

The robustness quantification method was applied to various phenotypic datasets to identify microbial strains with robust phenotypes, as well as to yeast deletion collections to investigate genetic markers of robust phenotypes. Additionally, it facilitated the exploration of robustness during laboratory evolution. The proposed method serves as a practical tool for various immediate applications, such as the selection of yeast strains exhibiting stable ethanol yields in beer production. In a broader perspective, grasping the mechanisms of robustness could facilitate the development of industrial strains.

1.1. Microbial robustness

Microbial robustness is an inherent characteristic of many microorganisms. In large-volume reactors typical of bioprocesses, microbial cultures must withstand widely varying gradients while maintaining consistent performance (3,4). To survive in dynamic environments or cope with internal perturbations, biological systems have evolved buffering mechanisms (1,2,5–8). Understanding robustness in biological systems helps not only to contextualize the observed trade-offs between different phenotypes (9), but also to design more resilient cell factories (10).

1.2. Perturbations and robustness quantification

An important aspect to consider when investigating robustness are the perturbations that prompted the emergence of robust mechanisms. In bioproduction, perturbations are associated mainly with variations in environmental factors, such as substrate composition, reactor gradients, and contamination (2). If perturbations are effectively replicated at laboratory scale, despite the many challenges related to such an endeavor (11), it becomes possible to quantify robustness and predict a strain's behavior on a larger scale. In this **thesis**, several quantification strategies were evaluated, with the mean normalized Fano

factor providing a reliable, precise, and standardized method. As highlighted above, quantification is used for initial strain selection and as a quality check for laboratory-designed strains before scale-up. Quantification and perturbation studies extend beyond bioproduction. Exploring the stability of fluorescent protein expression, for instance, proves valuable in synthetic biology (12); while investigating the robustness of microbial community composition provides insights on population dynamics across different contexts (13,14).

1.3. Intrinsic aspects of robustness

Although robustness can be studied from a phenotypic perspective, the underlying principles remain poorly understood (5,15). Proposed mechanisms include feedback regulation, redundancy, and modularity (1,16). Different approaches allow to explore the intracellular mechanisms of robustness, including rational genetic design based on tolerance studies (17) or mutant libraries grown under different perturbations (18,19). The quantification of robustness from phenotypic data and screens of yeast deletion collections performed within the scope of this **thesis** pointed to genes responsible for robust phenotypes. Even though the latter remain complex and difficult to describe in biological terms; various pathways and genes, such as the heat shock protein Hsp90 that regulates protein folding, have been suggested as key determinants of robustness (7,20).

1.4. Aim of the thesis

The primary aim of my thesis was to establish a method for quantifying microbial robustness. The secondary aim was to illustrate its application in examining robustness across phenotypic responses, as well as genetic and metabolic aspects.

Phenotypic data obtained from yeast cultivations in diverse environments were employed to test and validate the quantification formula, as well as to identify robust strains and their characteristics (**Papers II and III**). Evolution experiments using different methodologies revealed how robustness evolved over time and allowed to explore genetic hot spots through variant analysis (**Paper V**). The latter, combined with robustness quantification of variant mutant libraries established the basis for examining the metabolic and genetic aspects of microbial robustness (**Paper IV**) (Figure 1.1).

The following research questions were formulated in this **thesis**:

Which quantification theory is better suited to measure microbial robustness? Which experimental set-up can be used for this quantification?

To address these questions, I first cultivated yeast strains sampled from various sources in a perturbation set composed of nearly 30 different perturbations. The resulting dataset contained three dimensions: strains, perturbations, and measured phenotypes. Different formulae proposed by previous studies were tested on the collected data. After various iterations, the Fano factor proved to be the most suitable score (explanation in **Paper II**), in line with the definition of microbial robustness outlined in **Review Paper I**. The final formula for robustness quantification underwent adjustments to facilitate normalization and interpretation (**Paper II**).

How can the effect of perturbations be included in a microbial robustness measure?

Perturbations play a crucial role in understanding robustness, and their nature is intricately tied to specific robustness features. This **thesis** explores how various types of perturbations (**Review Paper I**) impact robustness. To evaluate phenotypic responses under different perturbations, diverse phenotypes and their robustness were assessed (**Papers II and III**). This analysis addressed broader biological questions, such as the potential trade-off between performance and robustness, or whether robustness is a general property rather than specific to the defined set of perturbations. Additionally, I explored how different sets of perturbations affected the phenotypic responses of a laboratory *Saccharomyces cerevisiae* strain carrying various gene deletions (**Paper IV**).

How can adaptive laboratory evolution be used to improve robustness?

Microbial robustness is referred to as the stability of phenotypes in the face of different perturbations (**Review Paper I**). This definition led to the question of whether fluctuating environments can serve as a selective pressure in evolving robustness within a laboratory setting. To test the validity of this hypothesis, I subjected various *S. cerevisiae* strains to an evolutionary experiment under fluctuating conditions (**Paper V**).

How can genetic and metabolic markers of robustness be identified?

Microbial robustness has rarely been linked to genetic and metabolic processes. The challenge arises from the absence of a broadly applicable quantification method, such as the one developed in this **thesis** (**Paper II**), and the complexity of the underlying genetic and metabolic architecture. However, exploration of genetic variants in the evolution experiment (**Paper V**) and analysis of the phenotypic responses of a yeast deletion library (**Paper IV**) highlighted a few crucial genes and metabolic processes associated with changes in robustness.

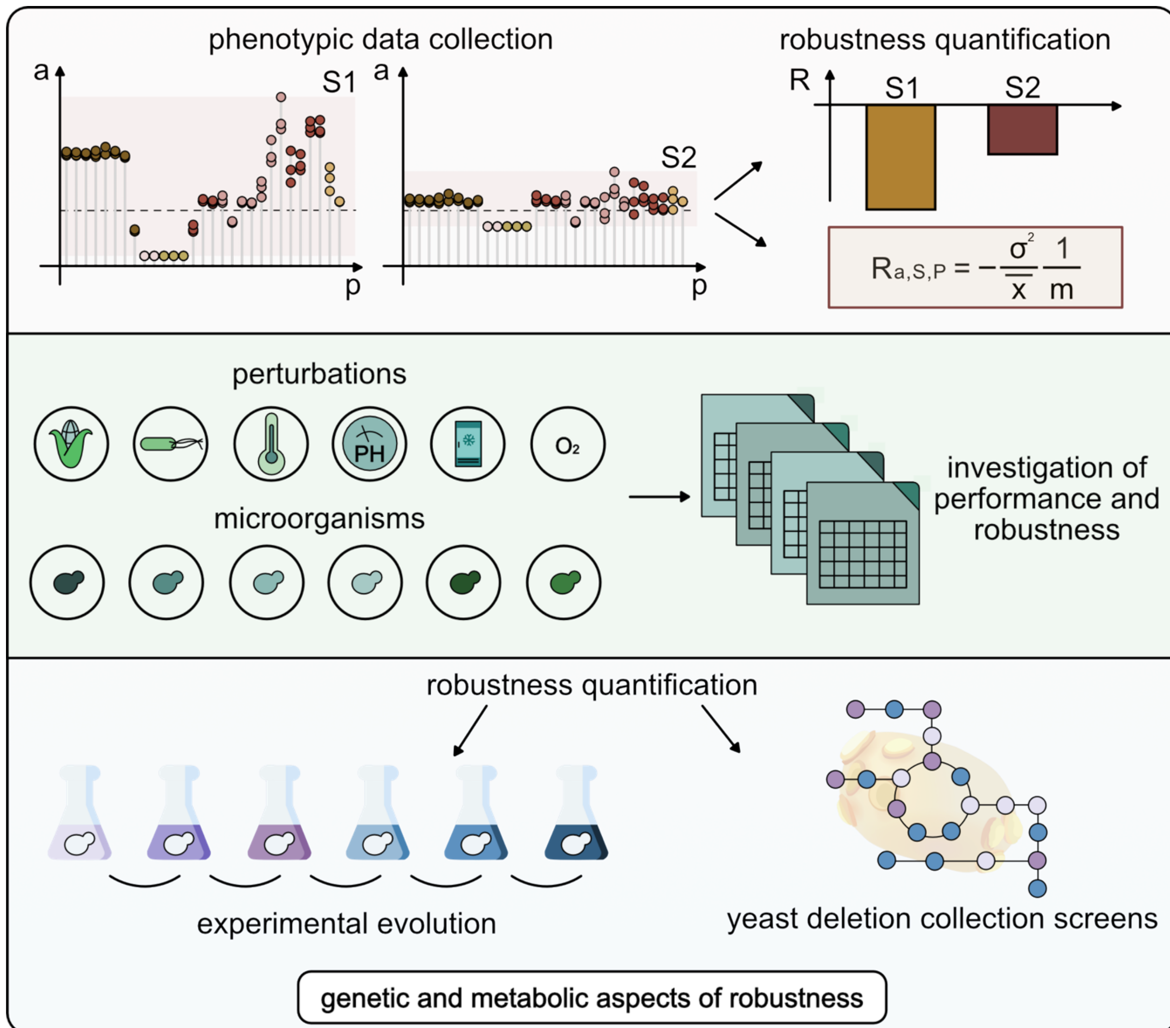


Figure 1.1: Overview of the research topics investigated in this thesis. *In this thesis a methodology to quantify microbial robustness was developed. Different aspects of robustness were investigated. Yeast cultivations in many environments as well as adaptive laboratory evolution were used in combination with the quantification method to characterize robustness in terms of phenotypic response. Quantification was applied to yeast deletion libraries and mutations from the evolution experiment were analysed to investigate robustness from a genetic and metabolic perspective.*

The long-term objective of my work is to contribute with tools and insights on molecular mechanisms of robustness that can be used to design robust cell factories and answer broader scientific questions.

In **Chapter 2**, the current potential of bioproduction and associated challenges are outlined. Special attention is given to the use of microbial cell factories within bioprocesses, the performance metrics currently employed, and the perturbations encountered in bioprocesses or beyond. Additionally, a brief overview of the strains, datasets, and phenotypic methods adopted in this **thesis** is provided.

Chapter 3 is entirely dedicated to the theoretical foundations and quantification of microbial robustness. It begins by presenting the definition and significance of robustness in various contexts. The work in this **thesis** is then contextualized within other studies on robustness, encompassing both theoretical and experimental investigations from different datasets. **Chapter 3** discusses the advantages and disadvantages of different quantification methods, including the one developed in this **thesis**. The perturbation space is placed in the context of robustness quantification, exploring combined effects of perturbations. The chapter concludes with a discussion on trade-offs, adopting a Pareto front perspective.

Chapter 4 explores robustness from a genomic and metabolic standpoint. My work on mutant libraries and evolution experiments is put into perspective with state-of-the-art knowledge. **Chapter 4** also explores applications of the quantification formula in other yeast deletion collections screens.

Chapter 5 provides a conclusive statement and explores future perspectives.

Chapter 2. Microbial Performance and Perturbations in Bioprocesses

The first part of **Chapter 2** is centered on the use of cell factories within bioprocesses and the metrics for evaluating their performance. Special emphasis is given to the yeast *S. cerevisiae*, which is featured throughout this **thesis**. The latter part is dedicated to perturbations that occur frequently during lignocellulose fermentation to produce ethanol.

2.1 Sustainable production of commodities

As countries try to address the United Nations' Sustainable Development Goals (SDGs), the integration of biotechnology and bioprocesses has emerged as efficient means to achieve prosperity, sustainability, and development. This **thesis** is driven by the objectives of SDG 13, which calls for urgent action to fight climate change and its effects. Specifically, this work can contribute to a sustainable bio-production that lowers CO_2 emissions and replaces fossil-based chemicals. To this end, I propose different tools to investigate robustness, with which robust and more efficient cell factories could be designed.

Since the industrial revolution, humanity has reaped the benefits of consuming fossil fuels, such as oil, coal, and gas, to generate heat, power engines, and produce electricity (21). Upon combustion, fossil fuels emit CO_2, the primary catalyst of climate change and a significant contributor to air pollution (22). The main sources of CO_2 emissions include electricity and heat production, as well as the transportation and manufacturing industries (Figure 2.1.a) (23,24).

In response to the concerns raised by elevated CO_2 emissions and their impact on climate change, countries participating in the United Nations Paris Agreement have outlined strategies to limit carbon emissions and shift towards more sustainable forms of energy. Currently, 194 nations have presented nationally determined contributions, in which measures to address climate change are embedded within national policies (25). Some of these policies envision the transition to various forms of renewable energy, with hydropower currently accounting for more than half the output (over 4000 TWh as of 2022) (Figure 2.1.b) (26,27). In Europe and the United States, wind farms could make up a similar share (28).

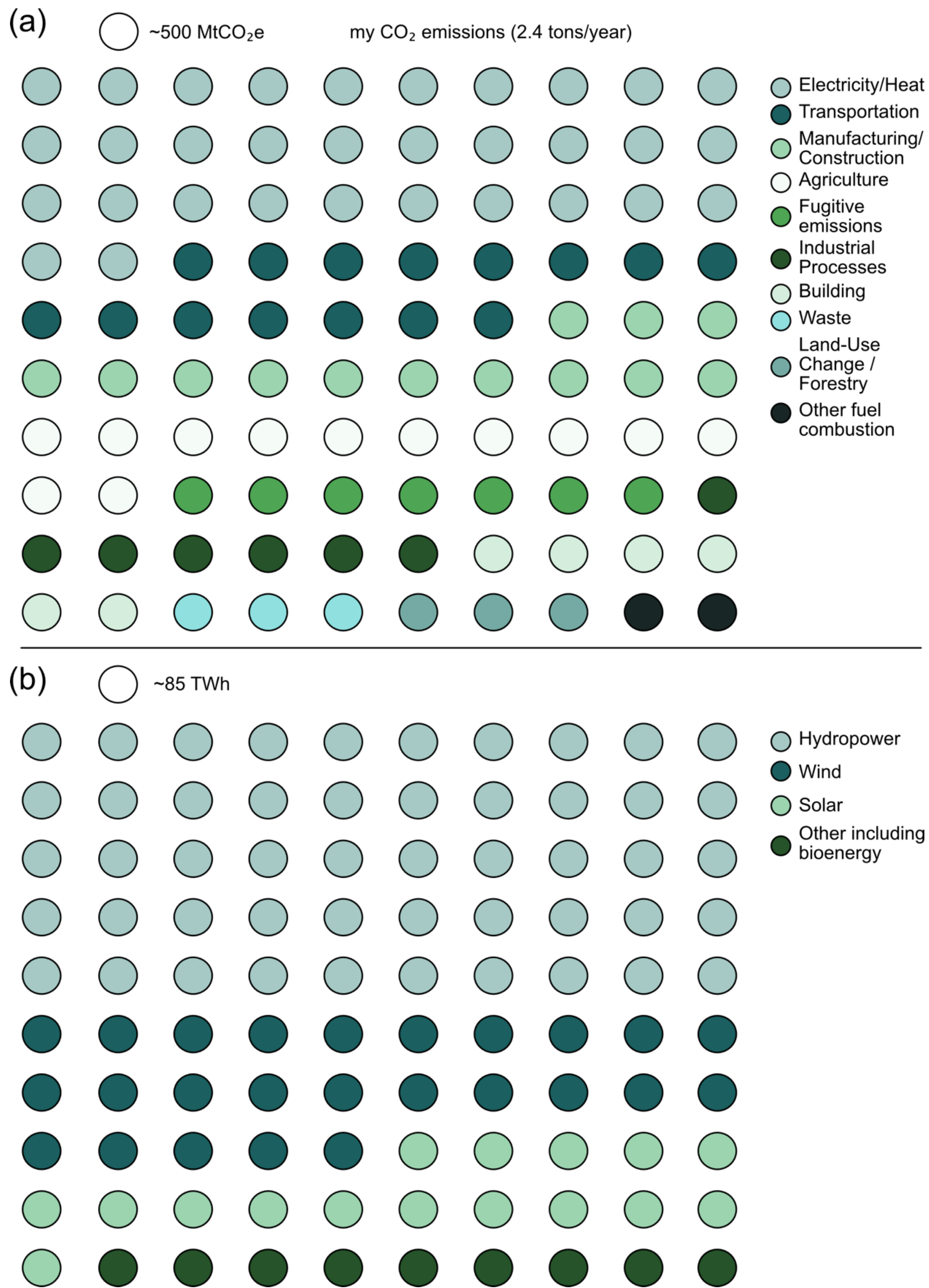


***Figure 2.1: World CO2 emissions and renewable energy generation**. a) Data were downloaded from (24) and normalized to total emissions. Data represents world emission in 2020. Each circle corresponds to approximately 500 MtCO2 equivalents and each sector is denoted by a different hue. b) Data were downloaded from (29) and normalized to total renewable energy production. Data represents world production in 2022. Each circle corresponds to approximately 85 TWh and each production source is denoted by a different hue.*

Another important source of renewable energy are biofuels. Derived from renewable feedstock, biofuels offer a sustainable alternative to fossil fuels. Ethanol, for example, is produced through the fermentation of renewable resources, such as corn or lignocellulosic biomass. It is commonly mixed with gasoline in different percentages for use in transportation. Bioethanol and biodiesel lower overall greenhouse gas emissions, provide a comparable net energy yield as gasoline, and are cost-competitive (30). Nevertheless, biofuels constitute a relatively small portion of the global renewable energy landscape. Only in the United States and Brazil, are biofuels and biodiesel derived from crops, such as sugarcane or corn mandated as components of transport fuel (31,32).

Achieving a more sustainable economy extends beyond bioenergy. Converting renewable resources into diverse bio-based chemicals could replace fossil-based counterparts and bioprocesses play a pivotal role in ensuring the success of this transition.

2.1.1 The requirement for robust biocatalysts

The success of bioprocesses relies on the effective execution of upstream (i.e., media preparation, cell inoculum), production (i.e., fermentation), and downstream (i.e., product purification, isolation, refinement) processes. Although all three phases are vital for achieving large-scale and high-quality output, rapid processing, and cost-efficiency, production stands at the core of a successful bioprocess (33). This **thesis** gives particular attention to the production phase and the microorganisms, also called biocatalysts or cell factories, involved in it. As these microorganisms should effectively convert the given substrates into desired products, production depends on the development of robust cell factories. This includes strategies aimed at minimizing by-products, balancing ATP and reducing power, and enhancing tolerance to industrial stresses, such as medium components, temperature, pH, and limited oxygen (4). Cell factory performance and industrial perturbations are described in detail in **sections 2.2** and **2.3** of this **thesis**.

2.2 Cell factories in bioprocesses

Cell factories capable of converting renewable resources into valuable chemicals are important players in the global economy (34). The European Commission has identified bio-based products that are already cheaper and environmentally friendlier than their fossil-based counterparts (35). Examples include the production of acrylic acid in *Escherichia coli* by BASF (Germany), Cargill (USA), and Novozymes (Denmark), as well as polyethylene by Braskem (Brazil) (36,37).

Model organisms, such as *S. cerevisiae* or *E. coli*, are widely used as microbial cell factories (38). The choice of these organisms is driven primarily by the availability of resources for their manipulation, including synthetic biology tools, sequenced genomes, and

comprehensive knowledge of their metabolism, as well as their rapid specific growth rates (39). Model organisms can be engineered to enhance the yield of primary or secondary metabolites of interest, biomass or macromolecules. Heterologous genes and pathways can also be introduced into the host to produce non-native compounds (40). While this approach comes with its benefits, there are various reasons why engineering a production host may not always yield the desired results. The integration of heterologous pathways or genes can often interfere with host metabolism. This interference may arise from the toxicity of the produced compound or poor understanding of regulatory mechanisms governing the heterologous pathway in the native and host microorganism, making it challenging to transfer these mechanisms into the host (41). Another potential hurdle stems from imbalances created by the diversion of resources and energy towards side reactions.

An alternative approach to using model hosts involves employing non-conventional organisms already possessing the desired characteristics and establishing large-scale production directly from them (42–44). Successful examples of this approach include the production of polyhydroxyalkanoates in *Clostridium sp*., or the synthesis of oleaginous compounds in *Rhodococcus sp*. (45–47). In such cases, there is a clear advantage in maintaining the metabolic pathway responsible for producing the desired product, along with all the associated regulatory mechanisms and post-translation protein modifications (48). Moreover, non-conventional organisms often exhibit remarkable levels of osmotolerance, thermotolerance or tolerance to inhibitors, thereby holding an advantage under the harsh conditions encountered during bioprocessing as compared to the laboratory (49). However, given that physiology of non-conventional hosts is less well known, it may be challenging to predict their behavior in various environments, particularly when scaled up. Enhancing yields or inhibiting by-products becomes more difficult in such cases, due to the limited availability of tools for metabolic engineering (48).

2.2.1 *Saccharomyces cerevisiae*

Owing to its central role in this **thesis** and its widespread use in bioprocesses, the yeast *S. cerevisiae* is described in more detail in this subchapter. Achieving optimal robustness and performance by cell factories hinges on a delicate balance between host physiology, product output, and tolerance to various chemical and physical stresses. *S. cerevisiae* stands out as a preferred choice in bioprocesses, primarily for its innate ability to generate ethanol and CO_2, which are valuable in the food industry (e.g., for beer production and baking) and as biofuels (50,51). The versatility of *S. cerevisiae* extends to its role in heterologous protein production, exemplified by the ground-breaking production of human insulin initiated by Novo (now Novo Nordisk) in 1987, which revolutionized diabetes treatment (52). As a model organism, *S. cerevisiae* is invaluable in exploring biological mechanisms, owing to its genetic similarity and conservation of essential processes with

more complex organisms. The ease of manipulation and well-established cultivation tools further enhance its suitability for investigating various biological processes (53). In the present **thesis** work, *S. cerevisiae* was used as a model to investigate differences in phenotypic responses to perturbations and robustness (**Papers II–V**).

S. cerevisiae comes in a multitude of strains, each uniquely adapted to its specific environment (54,55). Strains vary primarily in ploidy and genetic diversity, as demonstrated by a comprehensive genotypic/phenotypic analysis conducted by Liti and colleagues on 1011 isolates (**section 2.4:** Dataset 4) (56). The wide diversity among *S. cerevisiae* strains presents a remarkable opportunity to delve into species physiology and map the intricate relationship between genotypic diversity and phenotypic responses.

In **Paper V**, three specific *S. cerevisiae* strains—two laboratory: CEN.PK113-7D, S288C, and one industrial: Ethanol Red—revealed substantial differences in terms of single nucleotide polymorphisms, deletions, and ploidy. These genetic distinctions translated into visible phenotypic variations (**Paper V:** Figures S1, S2, and S3). Industrial strain isolates often exhibit higher tolerance to heat, low pH, and substrate inhibitors compared to laboratory strains such as *S. cerevisiae* S288C. They also have displayed increased fermentation rates compared to laboratory strains (57). In order to evaluate the differences in performance between laboratory and industrial strains, the laboratory strain CEN.PK113-7D and various industrial strains, including those isolated from cachaça by Brazilian distilleries (58) were cultivated in micro-aerobic conditions using microtiter plates and chemically defined medium containing 2% glucose (**Papers II and III:** Material and Methods; **section 2.4:** Dataset 1). Generally, the specific growth rate or ethanol yield of industrial strains was not consistently higher than those of laboratory strains (Figure 2.2.a). Moreover, no significant distinction in performance was observed among the three groups of strains (Figure 2.2.b) when uniform manifold approximation and projection was applied to the collected phenotypic data comprising five measured variables, 24 strains, and 29 conditions (**section 2.4:** Dataset 1). Hence, categorizing strains simply as laboratory or industrial could result in misleading generalizations. A categorization based on, for example, genetic differences or ploidy would be more accurate.

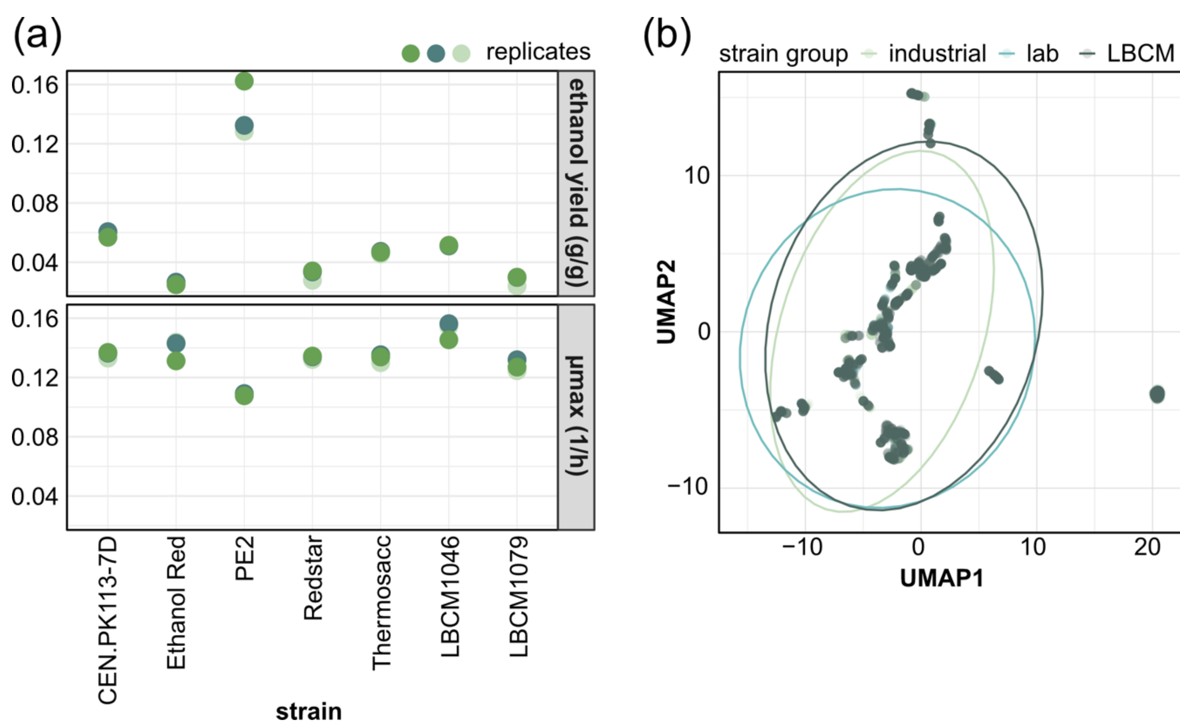


Figure 2.2: Performance comparison of S. cerevisiae strains. *a) Six industrial strains (two of which isolated from cachaça production sites in Brazil: LBCM1046 and LBCM1079) are compared to a laboratory strain (CEN.PK113-7D). Performance is shown for the maximum specific growth rate (µmax, 1/h) and ethanol yield (g/g). Different colors of the dots correspond to the replicates (n=3). B) Dimensionality reduction (uniform manifold approximation and projection: UMAP) was performed on phenotypic data from five measured variables measured under 29 conditions. The two principal dimensions are shown on the x and y axes. Each point corresponds to a strain and colored differently based on strain type (industrial: Ethanol Red, Thermosacc, Redstar, PE2; laboratory: S288C and CEN.PK113-7D; LBCM: LBCM collection from cachaça producing strains).*

Exploring the diverse phenotypic responses among *S. cerevisiae* strains is crucial (**Papers III** and **IV**), particularly when considering their potential application in industrial settings. Evaluating phenotypes, such as yields, titers, resistance to freeze-thaw cycles, shelf life, dehydration resistance or specific growth rates at different bioprocess stages, is important when assessing a strain's suitability for industrial application. However, even before testing potential microorganisms for production, it is essential to identify the right performance parameters to assess.

2.2.2 Performance metrics for the design of cell factories

When designing and testing cell factories, it is essential that performance measurements align with the specific objectives set for the process (59). Performance metrics serve to evaluate microbial performance and the corresponding economic feasibility of the process.

The following parameters are commonly included in the performance metric for benchmarking bioprocesses (60,61):

1) Titer or product concentration (mass of product/reaction volume)
2) Rate, productivity or space-time yield (mass of product/reaction volume/reaction time)
3) Yield (mass product/mass substrate)

Attaining elevated values in the titer-rate-yield (TRY) metric requires a concerted effort across various strategies (38). In particular, not all parameters in the TRY metric can be maximized simultaneously due to inherent trade-offs within microorganisms (62), as detailed in **section 3.3.1.**

TRY metrics directly linked to performance are crucial in designing cell factories, although they are not the sole parameters that demand attention. Numerous other factors are relevant in ensuring the success of bioprocesses. A universally applicable parameter for nearly all fermentation processes is the specific growth rate. On the one hand, simulations have suggested that a high specific growth rate can have a more significant impact than process productivity on strain selection (63). On the other hand, given that the biosynthesis of certain compounds competes with cell growth for energy and carbon resources, conducting fermentations at near-zero growth rates may minimize the diversion of feedstock towards biomass production (64).

Cell viability, especially in relation to cryopreservation, is rarely assessed. Cells are typically preserved in glycerol at -80°C. However, cryopreservation can potentially damage and diminish cell viability and overall survival, therefore making viability a potential factor in lowering the performance of the cells (65). The cell viability is a critical parameter in the production of bioethanol in Brazil, where the yeast pellet reused during fermentation undergo a viability check during the acid wash (66,67). The lag phase can pose challenges, particularly in batch and fed-batch fermentations. Typically, the lag phase is not a constraining factor in processes that use rich media as a substrate. However, when exploiting waste streams and less purified substrates, a long lag phase can become a limiting factor (68). A decline in productivity can be linked also to genetic instability and the gradual enrichment in non-producing cells over the course of fermentation (69–71).

It is beyond the scope of this **thesis** to investigate all these performance metrics and their impact; however, during the design of cell factories, it is important to keep in mind all aspects of a successful bioprocess. This can be facilitated by outlining the key design steps (Figure 2.3) and by engaging with experts in the field. In this **thesis**, I asked bioprocess experts to answer a survey comprising five questions related to bioprocesses and their associated challenges. I received five survey responses; two respondents had worked in

the bioprocessing sector for 5 to 10 years and the other three for less than 5 years. The participants held various positions, ranging from R&D to business development, and came from industries located across the globe. A combination of the design decision steps and survey answers is outlined below.

The first step of the decision process should define the goals of the bioprocess. The goals include evaluation of process outcomes, critical quality attributes, regulations in different countries where production is located, market value, costumers' demands, and environmental and societal impacts. Life-cycle assessment modelling tools can help define production goals (72).

The second step involves determining the relevant and measurable parameters that constitute the performance metric. This can be accomplished through analysis of comparable processes or a literature review (73). The objective is to look beyond the primary goal of desired product yield, and include often neglected secondary parameters such as cell viability. When surveyed about the performance metrics most commonly employed in bioprocesses, responders pointed to product yields, final titers, specific growth rates, cell survival and volumetric biomass productivity. These results underline how the choice of performance metrics varies significantly depending on the specific process being examined.

The identification of bottlenecks within a bioprocess allows for a more focused effort towards their elimination by optimizing performance parameters. Notably, most survey respondents identified the scaling up of microbial performance, coupled with downstream processing and fermentation-related capital expenditures, as significant bottleneck steps. The bottleneck arising from scaling up emphasizes the significance of my **thesis** work, as the methodology I developed to quantify the robustness of various cell factory parameters (discussed in **Chapter 3** and **Paper II**) could help predict, while still at a laboratory scale, strain performance in industrial environments.

The third and fourth steps should focus on the most suitable ranges and targets of the chosen metrics and their priority in relation to the objectives identified in the first step. Various tools can facilitate this process. Monte Carlo simulations for parameter sensitivity analysis and linearization techniques have proven effective for automating two critical quality by design tasks: i) assigning severity rankings for risk assessment and ii) formulating preliminary control strategies for critical process parameters (74). These methods have been successful in ranking parameters by importance and are essential given the inherent trade-offs encountered in bioprocesses, as discussed further in **section 3.3.1**.

The fifth step involves monitoring and controlling metric parameters. Fermentations are monitored with on-line sensors such as pH, temperature, and oxygen probes (75). Off-gas sensors are preferred for measuring CO_2, oxygen, and ethanol as they avoid interference with liquid and suspended particles (76). Vibrational Raman spectroscopy can

detect compounds in the fermentation media without the need for sample preparation. The implementation of biosensors, an effective synthetic biology tool, allows for the specific monitoring of compounds such as acetic acid (12,77). Infrared spectroscopy or biocalorimeters are employed to monitor biomass (78,79). Soft sensors and model-driven sensors play crucial roles in data analysis, process control, and optimization. Moreover, they have the capability to interpret data from other sensors and conduct fault detection (80).

In the survey, a question was posed about whether bioprocesses required more robust strains or more controlled processes. Despite significant progress in process control and sensor technology, all participants agreed that bioprocesses would benefit more from robust strains. They justified their choice by noting that highly controlled processes tended to be labor-intensive and expensive. If costs were equivalent, a combination of robust strains with controlled processes would be the preferred option.

In the sixth and last step, the process is optimized using observations from the previous steps. Here, data analysis is fundamental in the reporting and assessment of collected values, especially those from step five.

The decision process described above (Figure 2.3) can be applied to all aspects of a bioprocess, including fermentation technology and design of cell factories. In this **thesis**, particular attention was given to steps two to four, in relation to cell factory design and optimization.

Data from the cultivation of *S. cerevisiae* strains under 29 different conditions (**Paper III**, **section 2.4:** Dataset 1) showed that strains exhibiting a consistent product yield might also be characterized by slow specific growth rates or long lag phase. For example, the PE2 strain utilized in bioethanol production from sugarcane, displayed an exceptional high ethanol yield; however, its specific growth rate ranked the lowest among the scrutinized strains (Figure 2.4). In contrast, LBCM1079, with a specific growth rate comparable to the other investigated strains, demonstrated notably poor yields. Such observations highlight not only the natural trade-offs between performance parameters (discussed further in **section 3.3.1**), but emphasize also the need for a thorough decision-making process when selecting metric parameters to prevent scale-up delays and cost increases. Furthermore, a distinction among performance parameters is fundamental in the investigation of robustness (**Chapter 3**).

Finally, the environmental context in which performance metrics are assessed should also be considered. This implies a comprehensive examination of the industrial settings, with accompanying potential perturbations. Such approach is particularly relevant, if the ultimate objective is to develop a robust cell factory capable of maintaining consistent performance despite experiencing various perturbations.

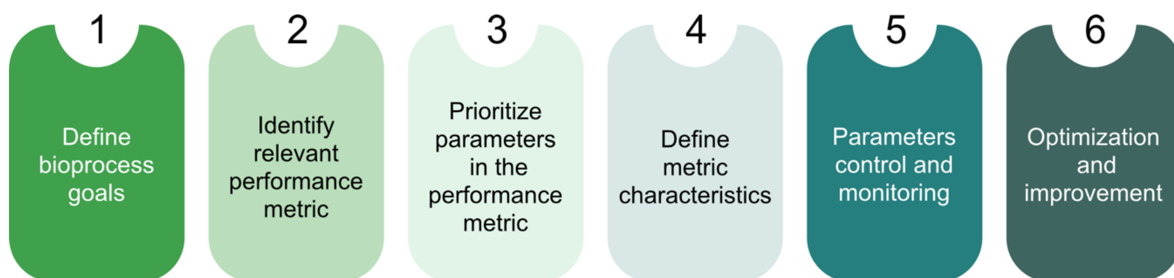


Figure 2.3: Bioprocess decision-making steps in relation to performance metric. *Illustration of the decision-making process in bioprocess design. Points 1 to 6 are described in detail above.*

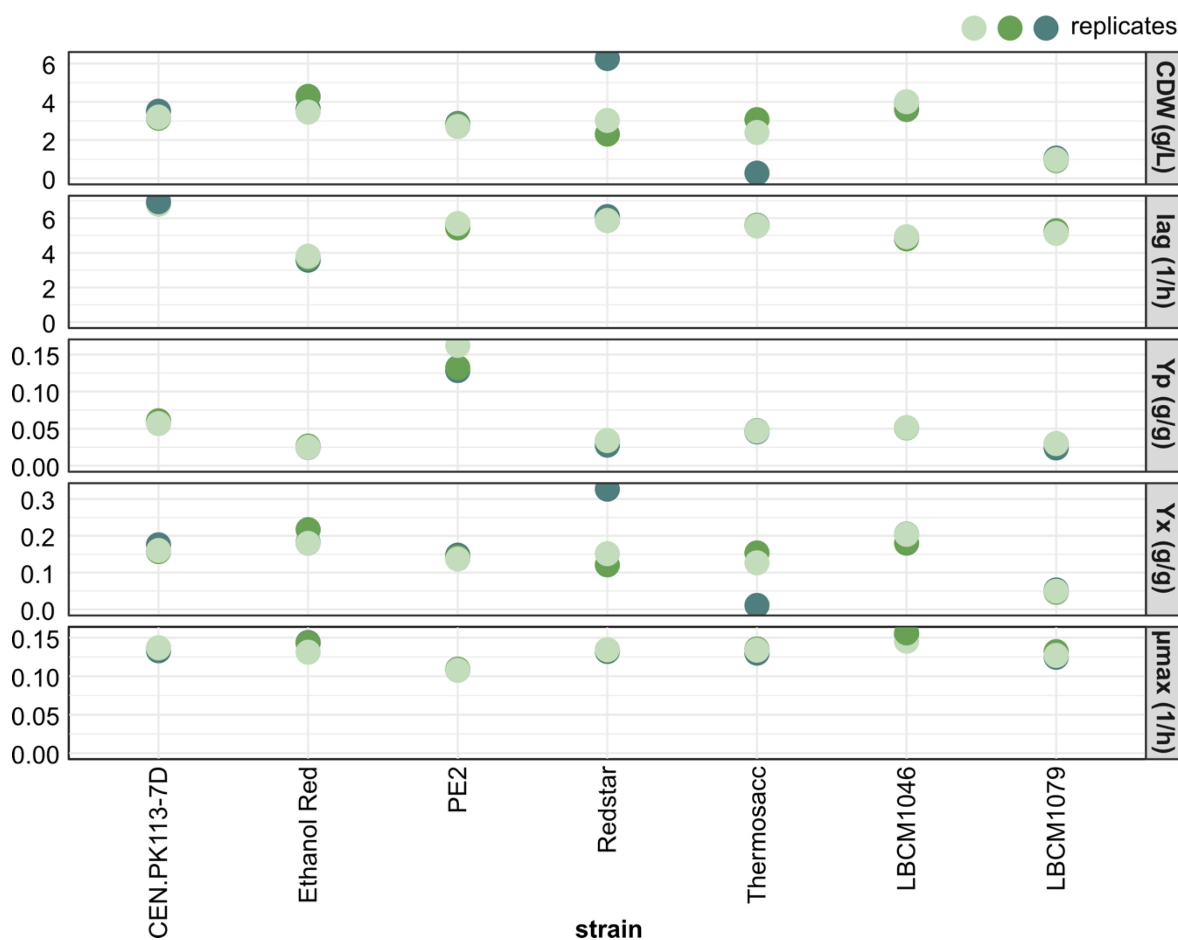


Figure 2.4: Investigation of different performance parameters in S. cerevisiae strains. *Six industrial strains (including cachaça strains LABCM1046 and LBCM79) and one laboratory strain (CEN.PK113-7D) are shown. Each dot corresponds to a replicate (n=3). Six parameters were measured: specific growth rate (µmax), lag phase (lag), cell dry weight (CDW), ethanol yield (Yp), and biomass yield (Yx).*

2.3 Perturbations

Perturbation is defined as an event that triggers a change in a system. A major challenge encountered when scaling up processes with laboratory-designed microorganisms is the identification and replication of large-scale perturbations in small-scale set-ups (61). Perturbations play a crucial role not only in bioprocesses, but also when studying cell physiology. In ecology, disturbances or perturbations are fundamental for species diversity and, ultimately, evolution. A strategy to improve predictions and streamline the scale-up process involves identifying potential perturbations and utilizing high-throughput testing to assess performance across a broad array of perturbations, referred to as the perturbation space. High-throughput testing has proven effective in uncovering mechanisms that were inaccessible through traditional rational approaches and in identifying strains with desired characteristics (81).

2.3.1 Perturbations in lignocellulose fermentation processes

In this **thesis**, the impact of perturbations on the fermentation of lignocellulosic biomass into ethanol was assessed in multiple papers (**Papers II–V**). Lignocellulosic biomass fermentation is a particularly relevant example owing to its various types of perturbations, relevance to the United Nations' SDGs, and abundance. The production of bioethanol from sugarcane, corn or sugar beet (classified as first-generation or 1G) started some 50 years ago and is now a well-established process. Bioethanol production from 1G feedstocks allows the United States and Brazil to replace part of their fossil fuels with biofuels. Instead, biomass-to-ethanol processes using second-generation (2G) feedstock, including corn stover, wheat straw or spruce, are marred by low cost-effectiveness and poor scalability (82). 2G biomass has been suggested as a more sustainable alternative to 1G sources, owing to concerns related to food security and land use (83). Agricultural and forestry residues, which constitute 2G biomass, are primarily composed of cellulose (35%–50%), hemicellulose (20%–35%), and lignin (84). The structural arrangement and proportion of these components depend largely on the plant type, season, country, and weather conditions. Such heterogeneity influences the accessibility of cellulose to hydrolytic enzymes and the subsequent release of fermentable sugars (Table 2.1) (85). Crystallinity, particle size, and accessible surface area play a crucial role in the success of enzymatic hydrolysis.

Table 2.1 Composition of lignocellulose biomass from different sources (adapted from (86))

Lignocellulose biomass type	Cellulose (%)	Hemicellulose (%)	Lignin (%)
Corn stalks	50	20	30
Sugarcane tops	43	27	17
Sugarcane bagasse	35	35.8	16.1
Corn stover	38.4	22.9	20.1
Rice husk	37.1	29.4	24.1
Rice straw	35.8	21.5	24.4
Spruce wood	43	29.4	27.6
Beech wood	44.2	33.5	21.8

To enhance enzyme accessibility, lignocellulosic biomass typically undergoes pre-treatment, via physical or chemical methods. Physical pre-treatments encompass milling, extrusion or microwaving; whereas chemical pre-treatments include acid, alkali or ammonia steam-explosion (86). While these pre-treatments are essential for increasing sugar accessibility, they also result in the release of inhibitory compounds that retard the fermentation process (Figure 2.5) (87). Hemicellulose hydrolysis yields pentoses and uronic acids, which undergo dehydration to form furfural; while hexoses are dehydrated to 5-hydroxymethylfurfural (HMF) (88). At elevated temperatures and acid concentrations, HMF can be further degraded into levulinic and formic acids, along with furfural. Additionally, acetyl groups from hemicellulose are hydrolyzed into acetic acid (89,90), and vanillin is released through lignin oxidation (91,92). These compounds can be present in relatively high concentrations, posing challenges to the fermentation process (87).

Following pre-treatment and enzymatic hydrolysis, the resulting lignocellulosic hydrolysate is utilized for fermentation. However, the presence of inhibitory compounds in this hydrolysate limits microbial performance (93–95). Weak acids exist in equilibrium between their dissociated and undissociated forms, with the shift in equilibrium depending on the dissociation constant (pKa) and the pH of the environment. At low pH, the undissociated form predominates and can diffuse through the cell membrane. Once inside, a slightly higher pH leads to its dissociation and consequent acidification. This event sets off a cascade of reactions. ATPase pumps protons out of the cell, diverting ATP from other metabolic processes such as biomass formation. Weak acids inhibit enzyme permeases, reducing the uptake of aromatic amino acids. Increased glycerol production implies a compromised capacity of cells to generate NAD^+; whereas an increased output of reactive

oxygen species can damage cell components. Finally, a lower DNA and RNA synthesis has also been observed (96).

Formic acid has been noted for its elevated toxicity in comparison to other acids (Figure 2.6). This increased toxicity is likely attributed to its lower pKa and smaller size, which facilitates its diffusion. In **Papers II and III** formic acid did not affect the cells significantly more than acetic acid, probably because at pH 5, most of the formic acid was in dissociated form. Prior research has indicated that formic acid leads to the downregulation of proteins associated with the biosynthesis and transformation of succinyl-CoA, as well as with alterations in the N-termini of core histones, thereby influencing cell development (97,98). While studies have indicated that levulinic acid leads to a more pronounced inhibition of cell growth (99), our findings show that levulinic acid had only a minimal effect on the specific growth rate (Figure 2.6). Lactic acid is not directly derived from lignocellulose pretreatment, but can be present in fermentations due to contamination from lactic acid bacteria (100). In my studies (**Papers II and III**), lactic acid had the lowest impact on fermentation (Figure 2.6), which could be attributed to its pKa and the pH during cultivations. In addition, lactic acid has been associated with an increased biosynthesis of glutathione (101). Interestingly, co-cultivation of *S. cerevisiae* with *Lactobacillus amylovorus* had been found to increase sugarcane fermentation yield by 3% when cross-feeding acetaldehyde, demonstrating a beneficial interaction between these two species (102).

Furfural and HMF inhibit microbial growth, even though they are converted to their less harmful alcohol counterparts under both aerobic and anaerobic conditions. HMF is thought to cause greater inhibition due to its slower conversion rate compared to furfural. Nevertheless, the reduction in fermentation and specific growth rate is more substantial with furfural than with HMF (103).

In the cultivations carried out in **Paper III** (**section 2.4:** Dataset 1), vanillin was found to affect strain growth more than HMF and furfural, even though *S. cerevisiae* can degrade it to vanillyl alcohol (104) (Figure 2.6). Vanillin disrupts the structural integrity of biological membranes, blocks translation by affecting the function of the large ribosomal subunit, and triggers the formation of processing bodies and stress granules. It also induces oxidative stress and mitochondrial fragmentation. The toxicity of vanillin is associated with chromatin remodeling, vesicle transport, and ergosterol biosynthesis (105).

High-throughput cultivations in conditions resembling the one encountered during lignocellulose fermentation (**section 2.4:** Dataset 1) provided insights on specific strain responses. For example, acid stress exerted a significantly higher impact on ethanol yield than on specific growth rate (**Paper III:** Figure 1). The same study showed poor acid tolerance by strains used in cachaça production (Figure 2.7). LBCM1001, LBCM1008, and LBCM1017 did not grow in the presence of acetic acid and LBCM110 failed to grow in the presence of almost all tested acids. Instead, most LBCM strains along with Ethanol Red

grew on 3 g/L furfural (Figure 2.7). Such specific behaviors point towards promising candidates for investigating the mechanisms underlying tolerance to different stressors. For instance, a phenotypic study comparing *Zygosaccharomyces bailii* with *S. cerevisiae* highlighted the former's pronounced tolerance to acetic acid. This disparity prompted further investigation into *Z. bailii*, whose enhanced acid tolerance was attributed to the structure and composition of its cell membrane (106,107).

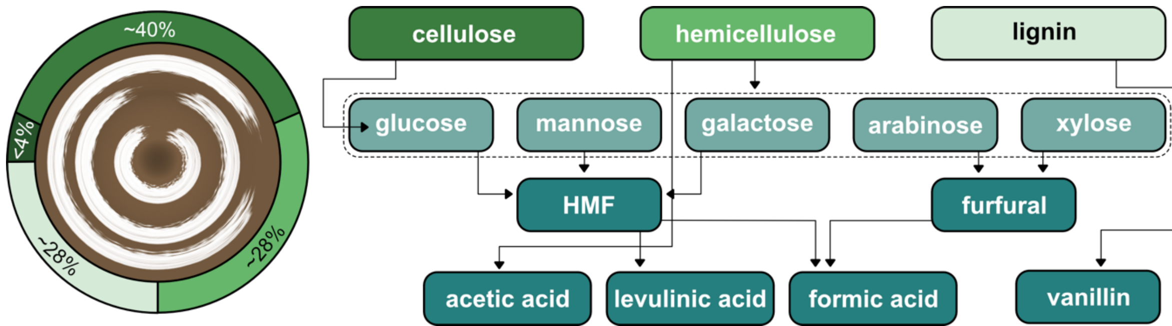


Figure 2.5: Composition of lignocellulose biomass. *Average percentages of hemicellulose, cellulose, and lignin are shown on the left. Sugars and inhibitors released during pretreatment of lignocellulosic biomass are listed on the right.*

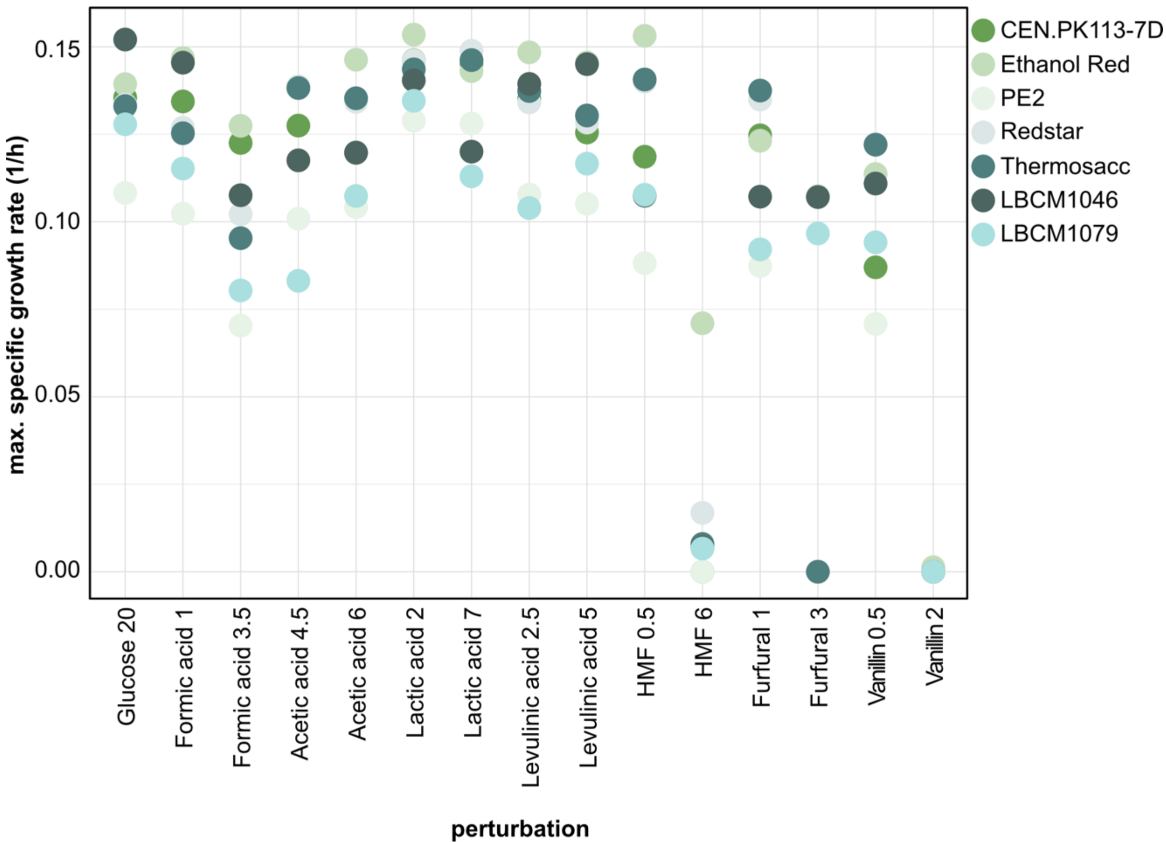


Figure 2.6: Specific growth rate of S. cerevisiae strains cultivated with acids and aldehydes. *Maximum specific growth rate is plotted on the y-axis and conditions are shown on the x-axis. Strains are plotted in different colors.*

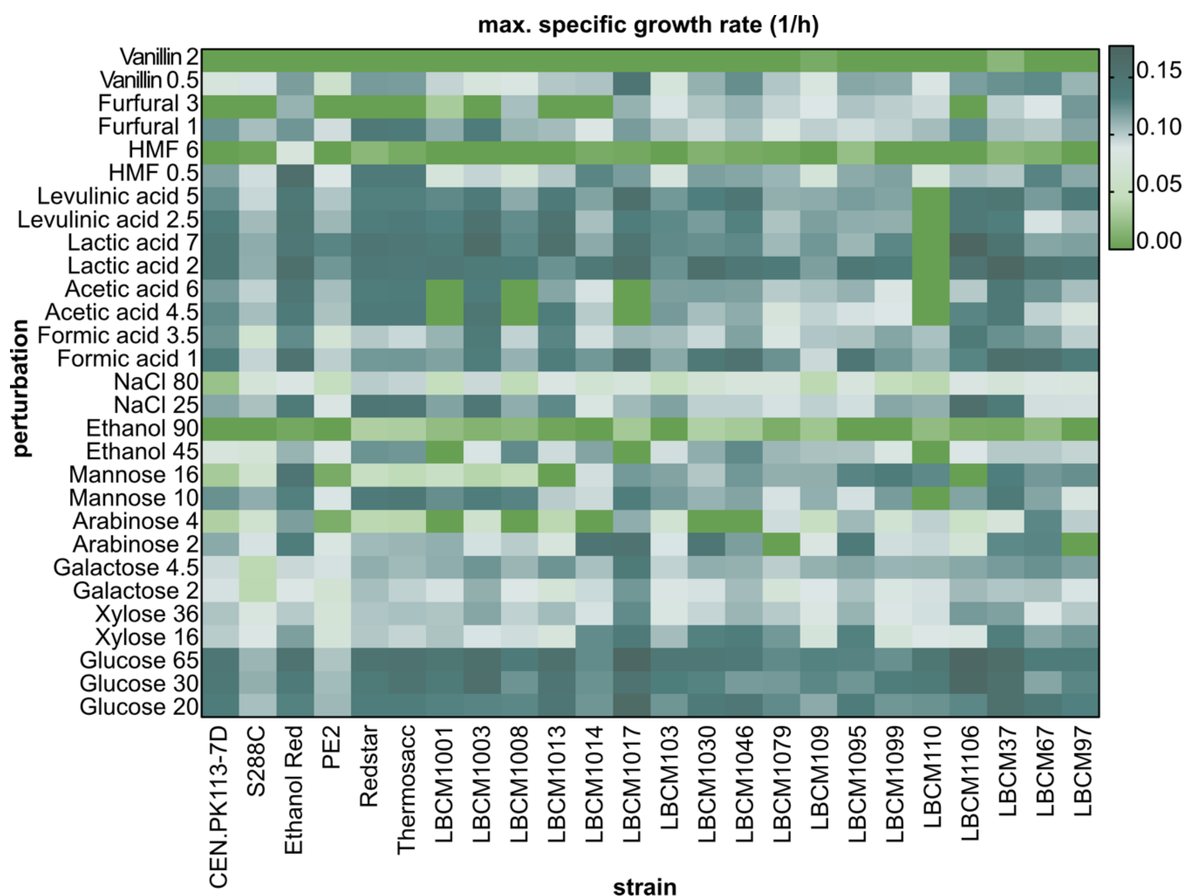


Figure 2.7: Specific growth rate of 24 Saccharomyces strains cultivated under 29 growth conditions*. Growth conditions are indicated on the y-axis and strains on the x-axis. Tile color represents the maximum specific growth rate of the strains.*

2.3.2 Bioprocess perturbations

The fermentation of lignocellulosic biomass into bioethanol serves as a notable example of how to effectively address complex processes by breaking down an intricate perturbation space into smaller, individual perturbations. The inhibitors present in the hydrolysates, along with sugars, are easily reproducible in controlled settings by preparing media with varying concentrations of inhibitors (**Papers II and III**). This approach finds application in another perturbation context, beer fermentation, as detailed in **Paper IV**. Within beer fermentation, yeast interacts with multiple complex substrates, impacting not only the fermentation outcome but also the taste profile of beer. Perturbations in this scenario did not exhibit a visible impact in the presence of a specific hop type or aroma. However, when compared to laboratory growth medium (Delft 2% glucose), the specific growth rate of both the industrial strain Ethanol Red and the laboratory strain CEN.PK113-7D was markedly lower (**Paper IV:** Figure 6). Once again, substrate variation proved to be straightforward to replicate compared to other types of perturbations such as pressure gradients. In the previously mentioned survey, when asked to identify significant perturbations in

bioprocesses, respondents answered that pH and temperature control, as well as reactor gradients, were the most critical factors affecting their processes.

Substrate variation represents only a fraction of the comprehensive array of perturbations faced by cell factories within industrial fermenters. Perturbations come in two forms: predictable and stochastic. Predictable perturbations are more manageable through process technology; whereas stochastic perturbations are more difficult to control and identify (**Review Paper I**). Anticipated perturbations encompass variations in substrate batches, inhibition stemming from both substrate and end products, gradients in nutrient and oxygen levels within large-volume reactors, and shear pressure (108). Cell performance is affected also by pH and temperature fluctuations. Instead, stochastic perturbations involve unpredictable events, such as declines in cell viability, contaminations, population heterogeneity, and genetic instability (109).

2.3.3 Three aspects in perturbation studies important for performance assessment

When assessing microbial performance under different perturbations, three aspects should be included:

I. perturbation probability, i.e., the probability that a certain perturbation will occur within the process.

II. perturbation intensity, i.e., the extent or magnitude of a particular perturbation.

III. interaction effects between perturbations.

The first factor concerns the relevance of a perturbation in relation to a specific process. An advantage of evaluating microbial performance across numerous perturbations lies in the ability to discern which ones are not pertinent to the process outcome. In beer production, assessing the performance of yeast strains with respect to furfural may be much less relevant than testing them against high levels of sugars or ethanol, as these conditions are predominant during brewing. Relevance of a perturbation to a process, known also as “perturbation probability”, has been estimated in the context of robustness assessment (15), but experimental quantification is not routinely performed. The challenge lies in attributing a probability/relevance value to each perturbation and estimating the likelihood of its occurrence in a process. Access to data regarding controlled parameters (e.g., temperature, pH or off-gas analysis) from industrial fermentations could inform such analysis.

To determine whether five phenotypes of 24 *S. cerevisiae* strains (**section 2.4:** Dataset 1) were influenced by the probability of occurrence assigned to each perturbation, random weights (sum of 29 weights equal to 1) were assigned to each of the 29 tested perturbations (unpublished data). Subsequently, the mean performance (for each of the five

phenotypes) of each strain across the weighted perturbations was computed (Figure 2.8). This process was repeated 1000 times with different randomly assigned weights. This resulted in 1000 mean values (across weighted perturbations) for each phenotype and strain. The distributions of calculated mean performances for all 1000 times were plotted for each phenotype and strain, as shown in Figure 2.8 for lag phase and cell dry weight. In the case of cell dry weight, only two strains exhibited a highly narrowed distribution (close to zero), suggesting that, regardless of the weight assigned to the perturbation, their performance remained remarkably stable. In contrast, other strains displayed wider distributions, indicating that perturbation probability had, in fact, an impact on performance evaluation. For the lag phase, strain distributions aligned into two distinct groups, each with similar modes. This alignment indicates that strain behavior was affected equally by different perturbation probabilities but with two distinct modalities. Ultimately, this test with randomly generated perturbation probabilities shows that certain strains are more influenced by the whole set of perturbations than others.

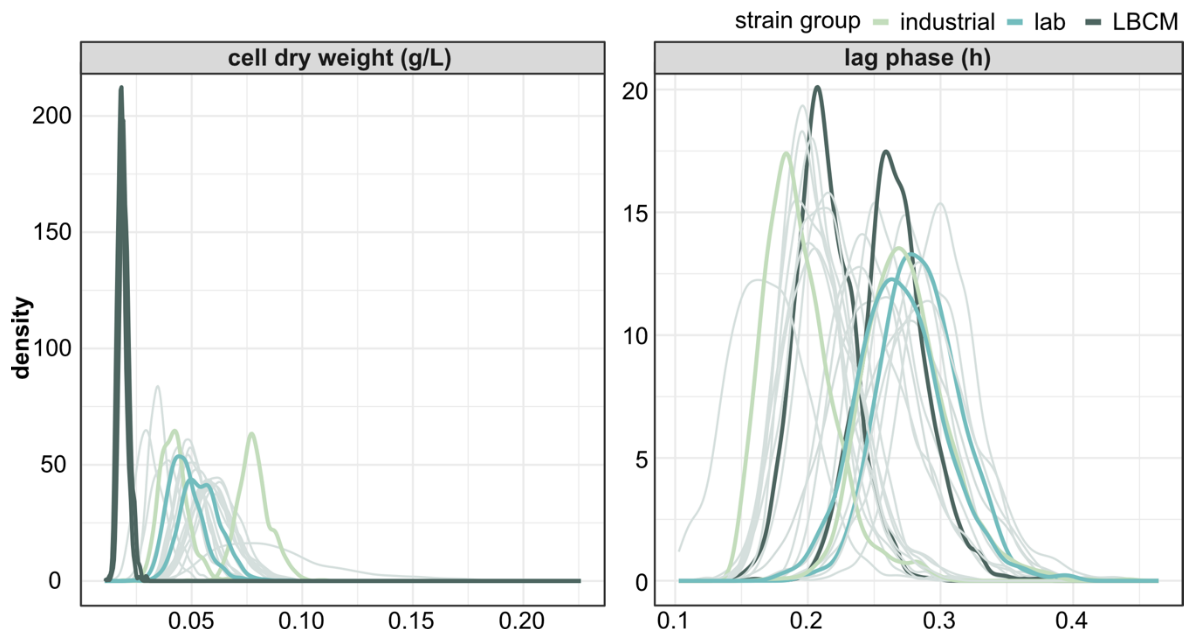


Figure 2.8: Distribution of phenotype weighted means over 1000 randomly assigned perturbation weights*. The distributions represent the weighted mean performance (random weight assigned to the perturbations) calculated 1000 times with different weights. Each color corresponds to a different strain type. The grey distributions belong to different strain types but just two strains per group were highlighted with a thicker line for simplicity.*

The second aspect involves perturbation intensity, meaning the amplitude of each perturbation (e.g., concentration of an inhibitor). If the concentration of a tested inhibitor is too low, it might not exert a discernible effect on the examined strain compared to when the perturbation is absent. Conversely, excessively high concentrations or perturbation intensities could severely impact cells, potentially inhibiting growth entirely. Consequently,

it is essential to conduct tests across a broad range of intensities. Lignocellulose hydrolysates vary in composition, depending on the type of biomass (**section 2.3.1:** Table 2.1), therefore sugars and inhibitors released during pretreatment can have very different concentrations. For instance, in the evaluation of *S. cerevisiae* strains PE2 and CEN.PK113-7D (unpublished data), four concentrations were tested for each perturbation (Delft media plus an inhibitor or different sugars) and are indicated after the name on the x-axis in Figure 2.9. The lower and higher concentrations were chosen based on published lignocellulose hydrolysate composition studies, offering a real although not necessarily optimal framework for testing. The specific growth rate of two exemplary strains was calculated for fourteen perturbation each divided in four concentrations (Figure 2.9). In the case of vanillin, HMF, and ethanol, the highest concentrations (3,9 and 123 g/L respectively) were excluded as no growth was detected. Subsequently, only two concentrations that partially blocked strain growth were chosen for further investigation (**Papers II and III**).

Finally, the third aspects that should be included when addressing bioprocess perturbations is the interaction effect of different perturbations occurring simultaneously. This arises from the concurrent presence of different perturbations, which may lead to synergistic interactions and downstream effects that are either more favorable or more detrimental than when the perturbations are evaluated individually. A study has found that high temperatures (37°C) lead to an increase in acetic acid production in *S. cerevisiae*, which, along with the presence of ethanol and reactive oxygen species, synergistically hinders cell growth and ethanol production rate (110,111). In the context of lignocellulosic biomass, the combination of liquid fraction of pretreated wheat straw mixed with ethanol and high temperatures resulted in no growth of two *S. cerevisiae* strains in a spot assay on agar (112).

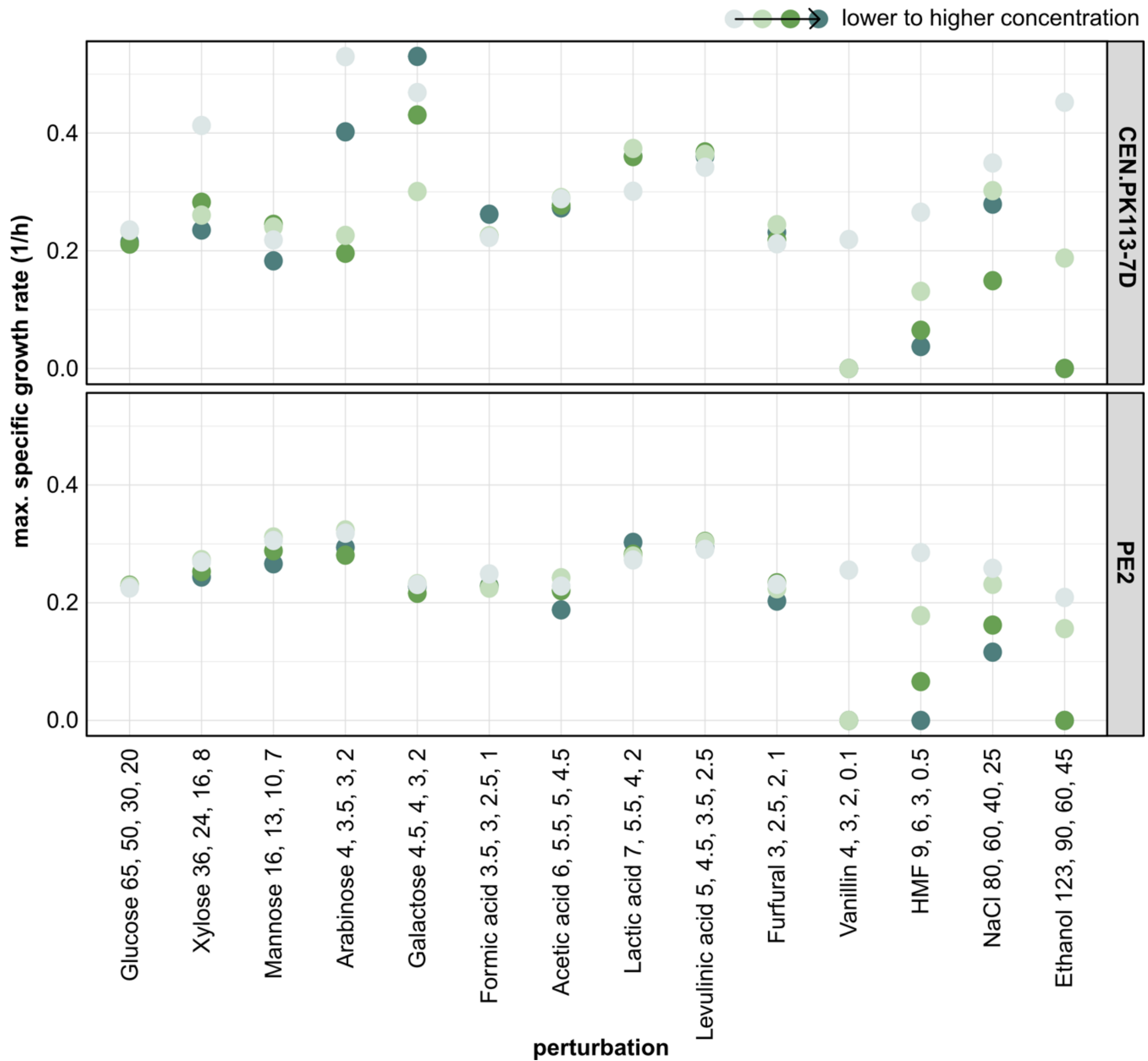


Figure 2.9: Maximum specific growth rate of two S. cerevisiae strains. *PE2 and CEN.PK113-7D were cultivated under 14 different conditions each containing the indicated compound at four concentrations (points of different hues). Concentrations of each compound are reported after the compound name in g/L on the x-axis labels.*

In this **thesis**, Ethanol Red and CEN.PK113-7D underwent testing under various combinations of inhibitors, and specific growth rates were compared to those of single inhibitors (Figure 2.10, unpublished data, **section 2.4:** Dataset 2). Notably, the specific growth rate for both strains exhibited a decrease when they were cultivated with a combination of two or three inhibitors. Combination of different acids with NaCl, aldehydes, and ethanol almost always resulted in no growth. The only condition that showed higher specific growth rate than the control (Delft medium + 2% glucose) was the combination of glucose (6.5%) with mannose (1.6%) and galactose (0.45%).

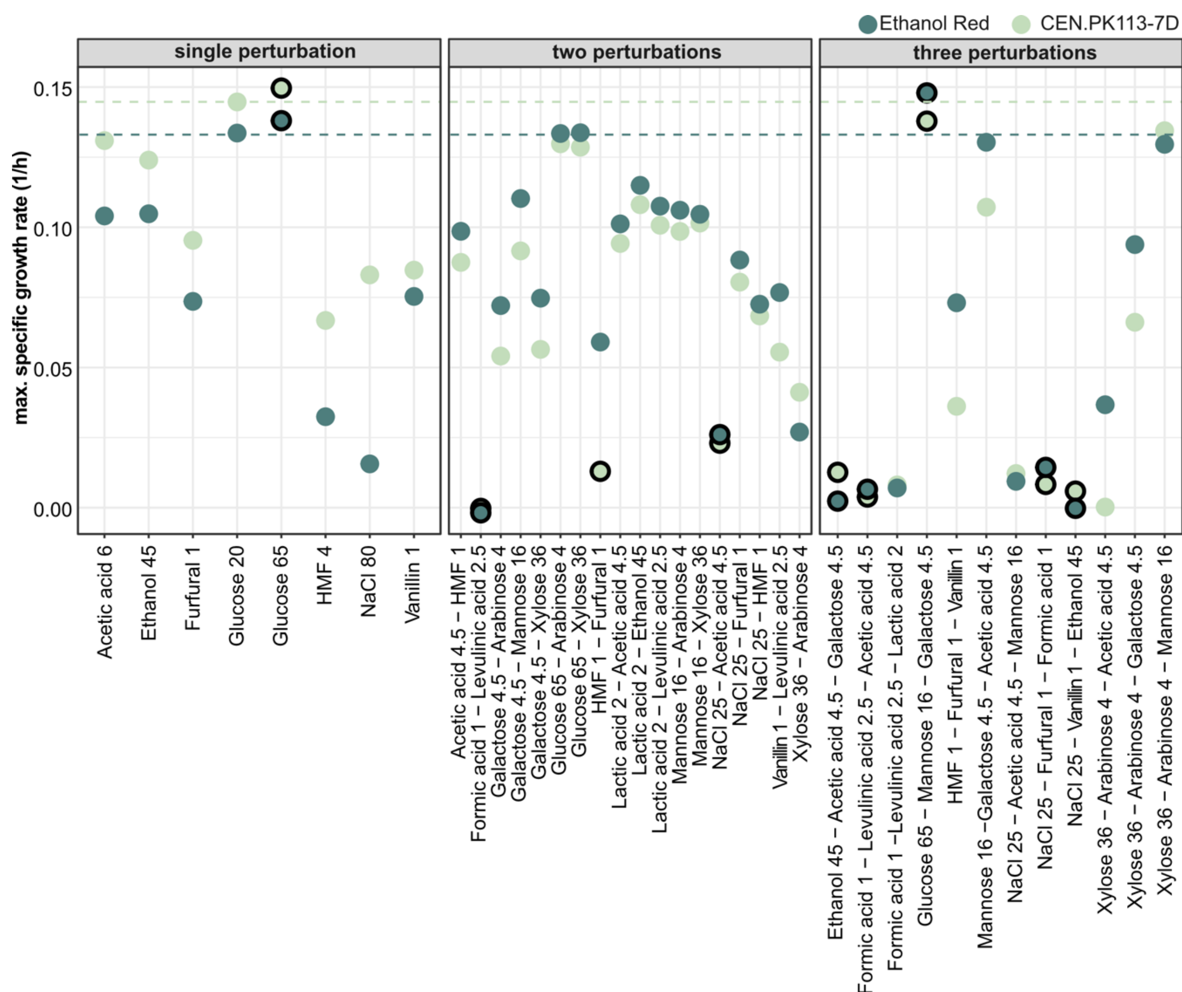


Figure 2.10: Combined effects of lignocellulose inhibitors on S. cerevisiae performance. *Maximum specific growth rate (y-axis) of CEN.PK113-7D and Ethanol Red cultivated under different conditions (x-axis). Different hues correspond to the two strains. Inhibitor name and concentrations are reported on the x-axis. Cultivations are divided based on the number of inhibitors tested simultaneously (up to three). The dotted line represents the control of both strains grown in Delft 2% glucose. The dots with a thick stroke line are highlighted as example and referred to in the text.*

2.4 Methods, datasets, and strains

In this **thesis** two main high-throughput methodologies were used to cultivate *S. cerevisiae* strains and monitor their growth: cultivation in 96-well plates using a growth profiler and cultivation on agar plates using the scan-o-matic. A detailed explanation of the methods can be found in **Papers II–IV** for the growth profiler and **Paper V** for the scan-o-matic; while a brief description is provided hereafter. The cultivation procedure employed in the evolution experiment involved a daily dilution in 96-well plates, and details can be found in **Paper V:** Material and Methods.

2

1. *Cultivation in 96-well plates using the growth profiler*

   Strains from a glycerol stock were thawed and 10 µL were transferred to 5 mL Delft medium to be cultivated at 30°C overnight. Optical density at 600 nm (OD600) of the overnight culture was measured and a volume corresponding to a starting OD600 of 0.02 was transferred to 96-well plates for a total culture volume of 250 µL. Multiple compounds were added to Delft medium to mimic different perturbations (e.g., acetic acid, vanillin). The 96-well plates were closed with either aerobic or micro-aerobic lids and placed in the growth profiler. Growth was monitored for 48 h at 30°C and 250 rpm shaking. The maximum specific growth rate and lag phase were calculated from the growth curves.

   To calculate cell dry weight and yields, the final OD600 of the culture was measured in a spectrophotometer and converted in cell dry weight through previously determined calibration curves. After removing the cells from the cultures, enzymatic assays were used to measure ethanol and sugars from the media at 48 h. The assays were based on enzymatic reactions, which produced NADH, whose absorbance at 340 nm was read in a spectrophotometer. Based on these measurements and initial concentrations, it was possible to calculate ethanol and biomass yields.

2. *Cultivation on agar plates using scan-o-matic*

   Strains from the glycerol stocks (kept on 96-well plates) were thawed on ice and pinned on a YPD agar plate in a 1536 colonies format. A normalization strain is generally pinned throughout the plate to adjust for variation. Colonies were pinned on a pre-culture medium (e.g., Delft medium in the case of **Paper V** phenotypic assays), after which the pre-cultures were pinned on agar plates containing the compound of interest (e.g., acids, ethanol). Colony growth was monitored in scanners for 72 h in four replicate plates to determine generation time and the produced biomass (calculated as final biomass - initial biomass) by the scan-o-matic software.

In this **thesis** work, I not only utilized datasets produced through my own experiments but also showed how applying robustness calculations to existing literature datasets could lead to new knowledge. The datasets used throughout this **thesis** are summarized below.

DATASET 1* (lignocellulose)	
Strains	CEN.PK113-7D, S288C, PE2, Ethanol Red, Thermosacc, RedStar, LBCM1001, LBCM1003, LBCM1008, LBCM1013, LBCM1014, LBCM1017, LBCM1030, LBCM1046, LBCM1079, LBCM1095, LBCM1099, LBCM1106, LBCM37, LBCM67, LBCM97, LBCM103, LBCM109, and LBCM110.
Applied method	Growth profiler + enzymatic assays
Perturbation space	29 conditions: glucose, xylose, galactose, arabinose, mannose, formic acid, acetic acid, levulinic acid, lactic acid, HMF, furfural, vanillin, ethanol, NaCl
Investigated phenotypes	Maximum specific growth rate, lag phase, ethanol yield, biomass yield, cell dry weight

*Details can be found in **Papers II and III:** Material and Methods

DATASET 2 (p. interactions)	
Strains	CEN.PK113-7D, Ethanol Red
Applied method	Growth profiler
Perturbation space*	Glucose, xylose, galactose, arabinose, mannose, formic acid, acetic acid, levulinic acid, lactic acid, HMF, furfural, vanillin, ethanol, NaCl
Investigated phenotypes	Max. specific growth rate

*Double and triple combinations were tested. For exact concentrations and combinations refer to Figure 2.10 (unpublished data)

DATASET 3* (Costanzo)	
Strains	4000 strains bearing single gene deletions (derived from *S. cerevisiae* BY4741 and BY7092)
Applied method	Solid growth
Perturbation space	YPD + 14 antifungal compounds
Investigated phenotypes	Normalized colony growth

*Specifics can be found in **Paper IV:** Material and Methods and (113)

DATASET 4* (Liti)	
Strains	1011 *S. cerevisiae* strains sampled from different locations and ecological origins
Applied method	Scan-o-matic
Perturbation space*	37 conditions (YPD + different compounds)
Investigated phenotypes	Colony growth normalized using growth on standard YPD medium at 30°C

*Details can be found in Material and Methods and Supplementary tables (56).

DATASET 5* (three p.spaces)	
Strains	Ethanol Red, CEN.PK113-7D and 14 CEN.PK113-7D strains carrying single gene deletions
Applied method	Growth profiler
Perturbation space	Lignocellulose hydrolysates perturbation space, Costanzo perturbation space, beer perturbation space each composed of 16 conditions
Investigated phenotypes	Maximum specific growth rate

* Details can be found in **Paper IV:** Material and Methods.

DATASET 6* (evolution)	
Strains	CEN.PK113-7D, S288C, Ethanol Red (1044 evolved samples, populations plus extracted single colonies)
Applied method	Scan-o-matic
Perturbation space*	20 conditions. Delft plus the following compounds: glucose, xylose, arabinose, mannose, formic acid, acetic acid, lactic acid, levulinic acid, HMF, ethanol, NaCl, no pH buffering, no trace metals, YPD
Investigated phenotypes	produced biomass

*Details can be found in **Paper V:** Material and Methods.

DATASET 7* (Persson)	
Strains	604 *S. cerevisiae* strains bearing single gene deletions
Applied method	Scan-o-matic
Perturbation space	Synthetic complete medium agar plates supplemented with 3 mM arsenite ([As III]; $NaAsO_2$), 4 mM arsenite, 0.25 mg/L rapamycin, 400 mg/L paraquat (methylviologen; N, N-dimethyl-4-4′-bipiridinium dichloride) or 1.25 M NaCl.
Investigated phenotypes	Produced biomass and doubling time

*Details can be found in Material and Methods (114)

Main Points from Chapter 2
• Bioprocesses are fundamental for a shift towards a more sustainable economy. There is a strong need for robust biocatalysts that can perform consistently in spite of perturbations.
• Performance metric is specific to the bioprocess and should include secondary parameters not only the TRY metric. Decision steps can help in identify and prioritize those parameters.
• Laboratory-designed strains should be tested for the decided performance metric before scale-up.
• Temperature and pH control, along with reactor gradients, were identified as major perturbations in bioprocesses.
• Perturbation probability, perturbation intensity, and interactions should be considered when testing strain performance in different conditions.

Chapter 3. Microbial Robustness

Chapter 3 describes the concept of microbial robustness and explores its significance across various contexts. It first describes the different methods for measuring robustness, with a focus on the Fano factor-based approach detailed in **Paper II**. This quantification technique is then applied to various datasets (described in **section 2.4**) to demonstrate its adaptability to different scenarios. The chapter concludes with a discussion on the trade-offs between performance parameters and between performance and robustness.

3.1. The concept of robustness

The term "robustness" has been defined in various contexts. Generally, it refers to the capability of a system to handle perturbations and maintain a stable output (115). This attribute is observed in various fields, such as aviation, where modern airplanes incorporate an automatic flight control system (AFCS) to maintain their flight trajectory. The AFCS detects deviations in the flight path caused by external perturbations and autonomously adjusts its input to restore stability. For instance, if strong winds cause an altered output, the AFCS will modify the route until the plane returns to a stable path (1). Another context is provided by machine learning, whereby robustness denotes the ability of an algorithm to produce consistent outputs even when exposed to noise or shifts in data distribution. It also implies that errors and properties present in the training datasets should align with those in the testing datasets to ensure stability (116)

In biology, metabolic robustness relates to the stability of diverse phenotypes, including regulatory mechanisms, gene expression, and phosphorylation (117,118). The ability to withstand genetic changes and mutations is referred to as mutational robustness, which is often linked to gene duplication (119). Transcriptional robustness relates to the ability of overcoming errors in transcription, splicing, translation, or post-translational modifications, and is crucial for preventing the production of non-functional proteins (120).

3.1.1 Microbial robustness definition

Within the scope of this **thesis**, most efforts have centered on microbial robustness as defined below and in **Review Paper I**.

Microbial robustness is defined as the ability of a system to maintain consistent performance despite external or internal perturbations (1,2,5,7,8).

Hence, by definition, microbial robustness is tied to three fundamental characteristics: the system (e.g., a strain, microorganism or population), the phenotype (and its measured

performance, also referred to as cellular function), and the perturbation space (**Paper II**). Phenotypes are manifestations of intricate sub features. Features, such as gene expression, enzymatic conversion rates, substrate uptake rates, and enzyme inhibition, can collectively influence ethanol yield (the main phenotype). In a simple scenario, if the phenotype "product yield" of a strain A exhibited two similar values when applying perturbations 1 and 2, then the phenotype "product yield" could be marked as robust across perturbations 1 and 2. Yet, the features of system A contributing to product yield when applying perturbations 1 and 2 may vary significantly (Figure 3.1). In fact, robustness tends to preserve functionality within a system despite perturbations, often demanding that the system dynamically adjusts its features (1).

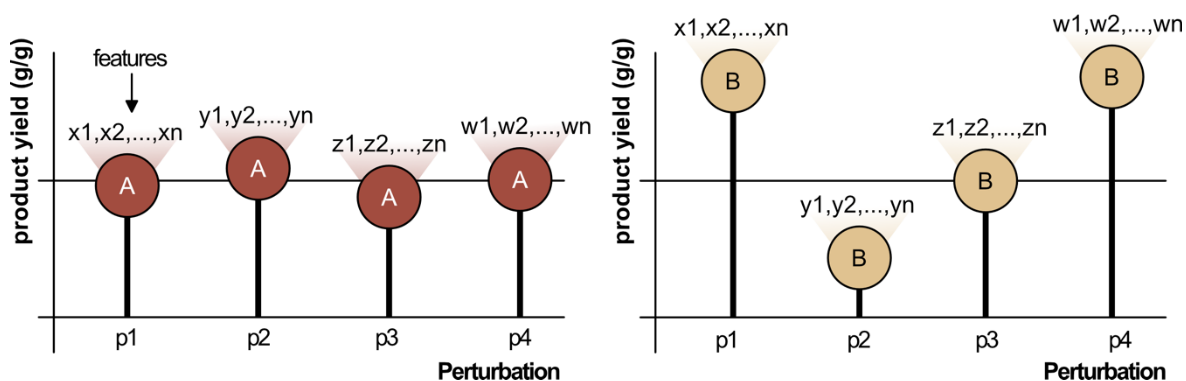


Figure 3.1: The robustness of phenotypes and its features. *The hypothetical product yield (y-axis) of systems A (red) and B (yellow) is represented in relation to four perturbations (x-axis). System A has a similar product yield in the four perturbations, which makes it is robust under the given conditions, even though it exhibits different features (x, y, z, and w). Conversely, system B, which also exhibits different features in the four perturbations, does not exhibit a similar product yield and, therefore, is less robust than system A.*

3.1.2 Relevance of microbial robustness

Microbial robustness becomes relevant when systems encounter perturbations and should maintain steady performance. While this quality is desirable in specific scenarios, it may assume a secondary role or be undesirable in other contexts. Bioprocesses aim for a stable and predictable outcome, particularly with respect to the TRY metric. In ethanol production from lignocellulosic biomass or side streams, robustness is indispensable to guarantee a consistent production flow, especially because of differences in substrate composition or batch-to-batch variations. In contrast, well-established processes such as insulin production may not rely so heavily on microbial robustness. The procedure has been meticulously fine-tuned over the past 40 years and employs highly pure sugars as substrates, thereby minimizing batch variations. The impact of other potential disturbances has been significantly mitigated through the optimization of strains and processes. In the United

States, Food and Drugs Administration regulations mandate "quality by design" in pharmaceutical manufacturing, with European countries following up on it. Process analytical technology supports and emphasizes the integration of quality into design, thereby reducing the focus on strain robustness (121).

As highlighted by the survey, there is a preference for designing robust strains over controlling processes (**section 2.2.2**), as well as for consistent performance over peak performance. Respondents pointed out that predictability and repeatability were essential for cost-effective operations. Additionally, the choice between robustness and high yield is influenced by product type. For products in high demand, the quantity produced (volumetric production rate) might outweigh the need for maximum yield; whereas titers, for example, become more critical for compounds that require extensive down-stream processing.

Robustness is an undesirable trait in cancer and antibiotic-resistant bacteria. Cancer is considered a robust system. While certain types, such as breast or ovarian cancer, exhibit a strong high initial response to chemotherapy, they often relapse, and the recurrent tumors frequently develop resistance to subsequent therapeutic interventions (122). Cancer is robust in the face of physiological challenges, such as low oxygen levels and metabolic stress. Extensive research over the years has linked this robustness to tumor heterogeneity, which arises from variability in size, morphology, antigen expression, membrane composition, proliferation rate, and metastatic potential (122). Heterogeneity offers a significant degree of modularity (different cells in the tumor are specialized in different tasks) and redundancy and is further strengthened through feedback controls, which contribute to robustness. The same can be observed in bioprocesses (**Review Paper I**) (123).

Another instance of highly robust systems is found in antibiotic-resistant bacteria such as *Pseudomonas aeruginosa*, which pose a significant threat to human health (124). *P. aeruginosa* exhibits a remarkable array of mechanisms to neutralize antibiotics, including low outer membrane permeability, efflux pumps expelling antibiotics, and the production of antibiotic-inactivating enzymes. Robustness is then further reinforced by the ability to form biofilms and horizontal transfer of resistance genes (125).

Understanding the mechanisms of robustness becomes crucial in such cases, as it enables the development of drugs that prevent disease recurrence after initial treatment.

3.2 Quantification of microbial robustness

To fully understand robustness and its inherent properties, it was necessary to develop a method for its quantification (**Paper II**). This would allow the evaluation of strain performance and comparative analyses of the targeted phenotypes. Performance evaluation provides insights on how effectively a system can navigate various perturbations and environmental conditions. For example, as elucidated in **Chapter 2** and detailed in

Paper III, the collection of performance data across various perturbations facilitates the identification of strains with a consistent behavior. Here, it enabled the assessment of performance across perturbations and whether it was consistently high or low for five distinct phenotypes. Ethanol Red was identified as a highly performing and robust strain (**Papers II and III**).

Robustness quantification can serve various purposes (Table 3.1). Engineered strains can be evaluated directly before scaling up or conducting quality checks for industrial strains (**Review Paper I**). Physiological studies can include measurements of the stability shown by intracellular parameters in different media. Furthermore, population heterogeneity can be quantified using microfluidic devices (126,127).

Table 3.1 Applications of the robustness quantification formula

Scope	System	Robustness calculated across:	Examples
Study trade-offs between phenotypes and between performance and robustness	Strains	Perturbations	**Paper III,** Dataset 1 (128)
Identify highly performing strains with robust phenotypes	Strains	Perturbations	**Paper III**, Dataset 1, Dataset 4 (56,128)
Identify genetic and metabolic markers of robustness	Strains bearing gene deletions	Perturbations	**Paper IV**, Dataset 3, Dataset 5, Dataset 7 (113,114)
Quantify robustness in phenotypic assays from evolution experiments	Evolving populations	Perturbations	**Paper V,** Daset 7 (114)
Evaluate the impact of perturbations on strains performance	Strains	Strains	(126)
Establish the degree of intracellular parameter fluctuations	Strains	Time	(126,127)
Quantify population heterogeneity	Single cells, subpopulations	Time, perturbations	(126,127)

3.2.1 State-of-the-art in robustness quantification

Various methodologies have been suggested to measure robustness (**Review Paper I**). Equations commonly employed in theoretical assessments of robustness include Kitano's formula (Equation 1) (15) and the one utilized by Yang et al. (129). Kitano's approach involves quantifying the robustness of a system "S" and a function "a" by integrating the evaluation function "D" across numerous perturbations "p" (perturbation space = P) multiplied by the single perturbation frequency "ψ(p)". The evaluation function represents the ratio of the function "a" in a perturbed state "p" with respect to the same function in a non-perturbed state "p0" (Figure 3.2.a).

$$R_{a,P}^{S} = \int_{P} \psi(p) D_{a}^{s}(p) dp \qquad \text{Equation 1}$$

Equation 1 was previously employed to conduct conditional robustness analysis and offer a valuable tool for exploring biochemical interaction networks in cancer (130,131). An alternative approach to assess robustness is through the coefficient of variation (CV), which is defined as the ratio of standard deviation to the mean of a measured quantity (129). Equation 2 defines robustness R as 1 minus the CV.

$$R = 1 - \frac{\sigma}{\mu} \qquad \text{Equation 2}$$

While these two equations have found various applications, they pose challenges when applied to the microbial robustness definition introduced in **section 3.1.1** (**Paper II**). Issues arise primarily from the selection of arbitrary reference conditions to determine perturbation probabilities, and the inherent difficulty of comparing robustness for different phenotypes because of scaling and unit differences.

3.2.2 Fano factor-based quantification

To address the above-mentioned challenges, a different formula (Equation 3) for quantifying robustness was introduced in **Paper II**. Equation 3 was derived based on prior research, in which the Fano factor had been proposed as a reliable method for measuring both robustness and variation in phenotypic data (7,132). Equation 3 measures robustness of a system (*S*) and a phenotype (*a*) across a perturbation space (*P*) by dividing the variance (σ^2) by the mean ($\bar{x}$) of the data and normalizing with the mean of performance data across all tested strains (m) (Figure 3.2.b).

$$R_{a,S,P} = -\frac{\sigma^2}{\bar{x}}\frac{1}{m} \qquad \text{Equation 3}$$

This formula has the following benefits (outlined in **Paper II**):

i) it eliminates the need for arbitrary control conditions and perturbation frequencies (unlike Kitano's formula, which requires p0 and ψ(p));

ii) the introduction of 1/m for normalization allows for the comparison of robustness across different phenotypes within the set of analyzed strains;

iii) the negative sign enables the representation of stronger robustness with higher values (with 0 signifying maximum robustness and no data variation).

In contrast to Equation 2, whose R becomes negative when the CV is >1, therefore, complicating interpretation, the Fano factor emerges as a more suitable representation, particularly for data approaching zero. This is crucial, considering the broad spectrum of values phenotypes can encompass (133).

When evaluating robustness quantification methods in this **thesis**, 1000 data points aligned along eight distinct distributions were generated using different functions in R softare (e.g., normal - *rnorm*, uniform - *runif*, and exponential - *rexp*) (Figure 3.3.a). Subsequently, Equations 1, 2, and 3 were employed to calculate the robustness of the generated data. To simplify the process, the p0 in Equation 1 was assigned a random value within the distribution range, and the probability of each perturbation was disregarded. Notably, modifications were made to Kitano's formula to adapt it for discrete data analysis. Multimodal and normal distributions exhibit a higher spread. Intuitively, based on the microbial robustness definition, one would anticipate lower robustness for the multimodal and normal distributions. However, Kitano's formula yielded unexpected results, associating exponential and binomial distributions with the lowest robustness values (Figure 3.3.b).

When employing the CV for robustness calculation, the multimodal and normal distributions attained the lowest values. Yet, R was negative for the multimodal distribution, which complicated the interpretation of robustness, because Equation 2 expects robustness to fall within the range of 0 to 1. Akin to the CV, Equation 3 assigned the lowest robustness to the multimodal and normal distributions. Comparable values emerged for the remaining distributions, with the gamma distribution achieving the highest robustness. Consequently, even when confronted with diverse data distributions, Equation 3 appears to be better at quantifying microbial robustness according to the four criteria outlined in **Paper II**: statistical significance, independence from performance, standardization, and comparison among phenotypes. Equation 3 can be applied also to a widely varying range and scale of values, as demonstrated when quantifying robustness for cell area or protein fluorescence (126,127).

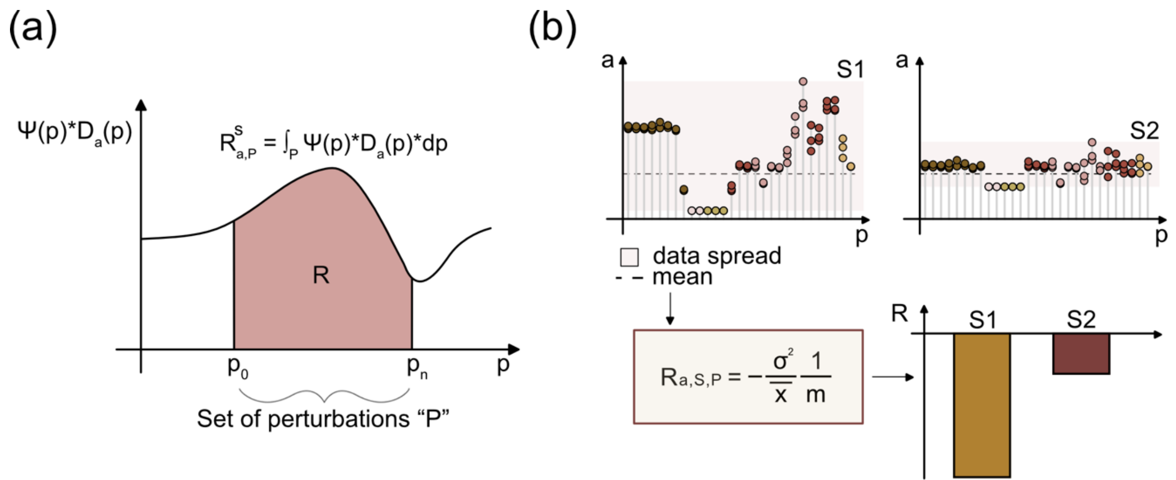


***Figure 3.2: Quantification of robustness**. a) Quantification of robustness according to Equation 1 (adapted from **Review Paper I**: Figure 5). The x-axis corresponds to the perturbation space and the y-axis to the evaluation function. b) Quantification of robustness according to Equation 3 (**Paper II**). The x-axis corresponds to the perturbations and the y-axis to the performance of phenotype "a". The dotted line is the mean and the colored rectangle to the spread of the data. An example for two hypothetical systems S1 and S2 is given. Equation 3 is shown in the box and the quantified robustness is represented by bar plots.*

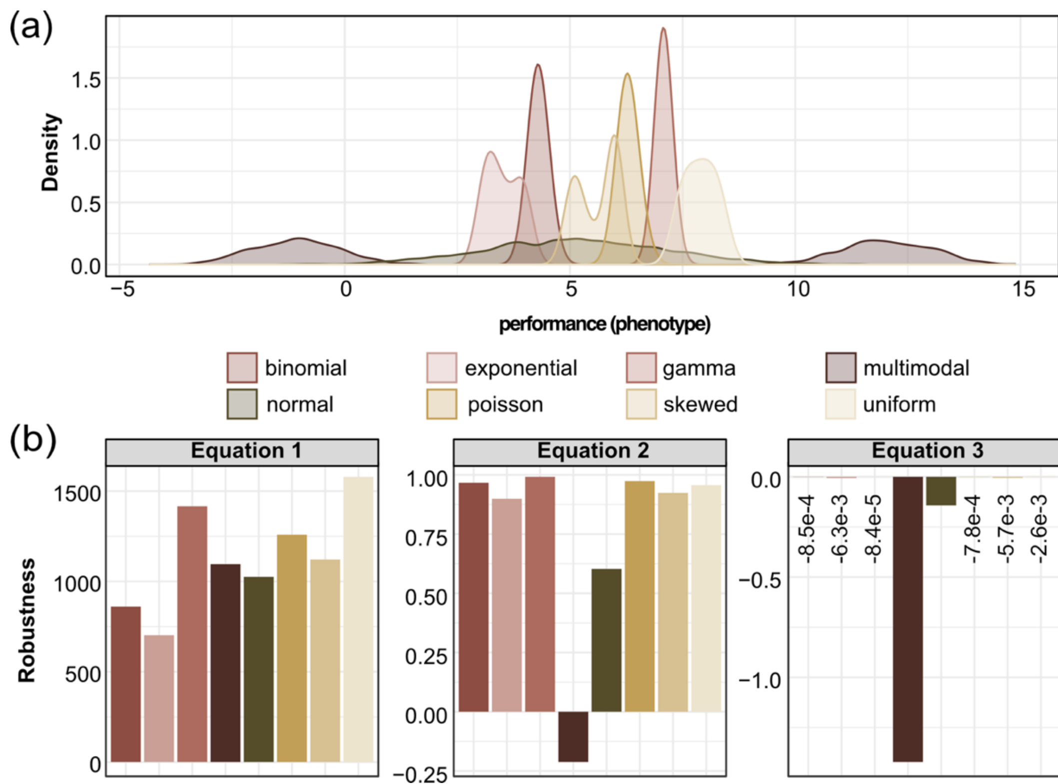


***Figure 3.3: Quantification of robustness with different formulae**. a) Eight different distributions (colors) are plotted on the y-axis with performance values of a hypothetical phenotype on the x-axis. b) Robustness of the generated data for each distribution was computed using Kitano's formula (Equation 1), Yang's formula (Equation 2), and the formula based on the Fano factor (Equation 3, **Paper II**). Robustness is reported on the y-axis and the different distributions are denoted by different colors.*

3.2.3 Robustness and tolerance

Exploration of robustness properties and definitions highlighted a discrepancy in terminology. Earlier studies often used the term "robustness" when discussing what would be more accurately described as "tolerance" (2). As elucidated in **Review Paper I**, tolerance is defined as the ability of microorganisms to survive in the presence of specific stressors. For example, tolerance to acetic acid refers to the maximum specific growth rate of strains in increasing concentrations of acetic acid (134). The concept of tolerance is also relevant in antibiotic resistance, whereby tolerant bacteria withstand increasing concentrations of antibiotics. This phenomenon was vividly demonstrated in a compelling experiment conducted by Baym and colleagues on a MEGA agar plate, on which bacteria were able to grow at increasing concentrations of antibiotics (135). In contrast, robustness does not specifically relate to survival and specific growth rate. Microbial robustness is indicative of the stability of performance across a broad spectrum of concentrations, perturbations, and phenotypes (**Review Paper I:** Figure 2).

Tolerance can be studied by measuring different phenotypes, including the specific growth rate in the presence of different concentrations of chemicals. In Figure 3.4.a depicts robustness for strains with similar or different specific growth rates and tolerance. The same data from **section 2.3.3** (specific growth rates of two *S. cerevisiae* strains CEN.PK113-7D and PE2 grown at four concentrations of different chemicals) were used to evaluate tolerance and robustness. Their maximum specific growth rates are shown in Figure 3.4.b for three different growth conditions and four concentrations of chemicals. Notably, when evaluating robustness across different concentrations of the same stressor, the specific growth rate of PE2 remained relatively consistent in the presence of acetic acid; whereas that of CEN.PK113-7D was stable in the presence of glucose alone. The specific growth rates of both strains decreased in the presence of HMF; although CEN.PK113-7D was more tolerant because it survived at higher concentrations of HMF than PE2. Robustness quantification offered a more general depiction of strain stability across concentrations for all fourteen tested conditions, something not possible with tolerance studies. In the above case, robustness of the specific growth rate calculated with Equation 3 revealed that P2 was more robust than CEN.PK113-7D (-0.16 vs -0.26 respectively).

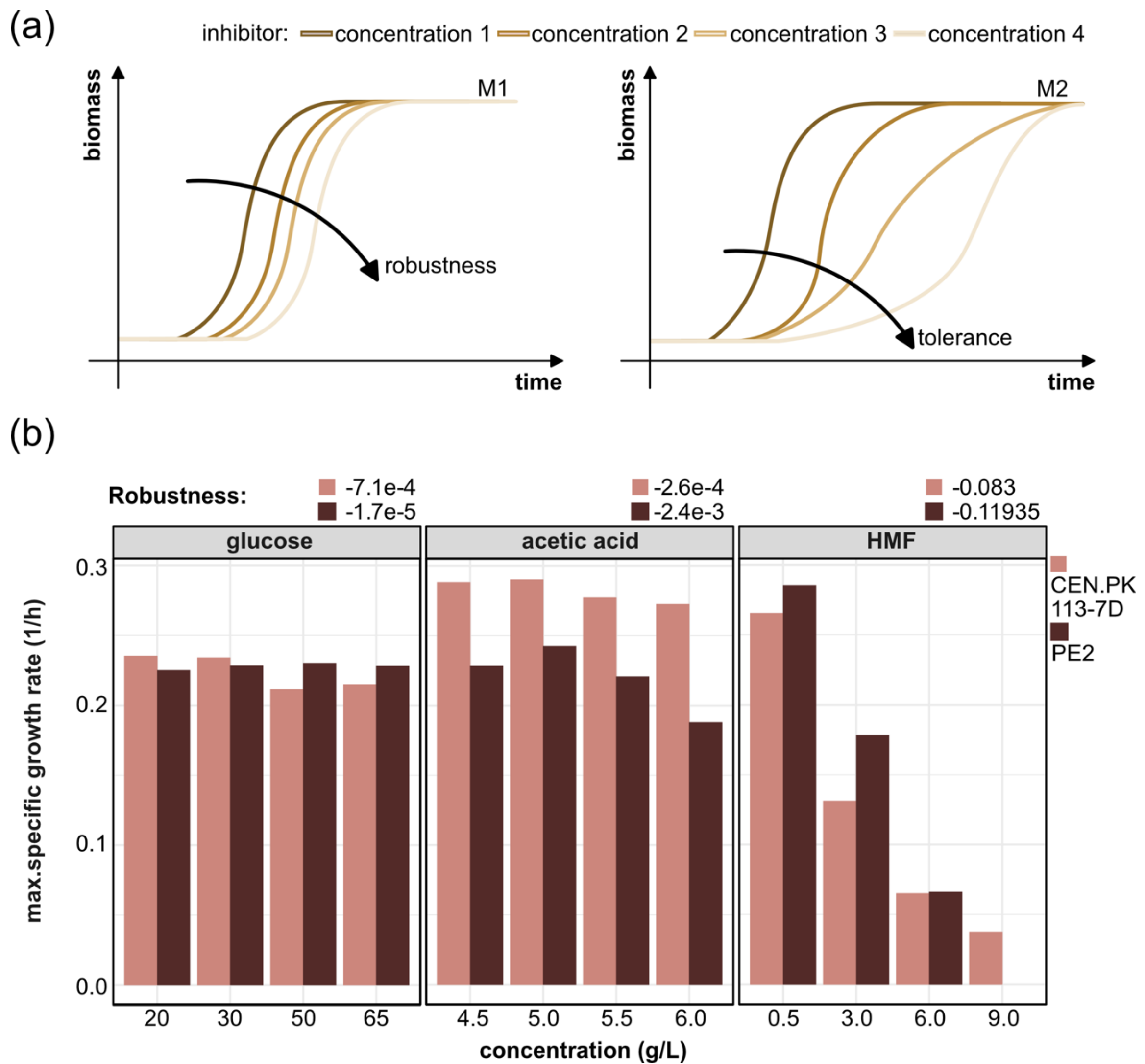


Figure 3.4: Robustness vs tolerance. *a) Growth curves (biomass vs time) refer to two different yeast strains, M1 and M2, cultivated in the presence four different concentrations of a specific inhibitor (different hues). The left graph illustrates robustness, with similar specific growth rates; the right graph showcases a less robust strain, tolerant to the inhibitor up to concentration 4. b) Maximum specific growth rate of two S. cerevisiae strains, CEN.PK113-7D (pink) and PE2 (brown)(y-axis) with respect to four different concentrations of three inhibitors (different facets)(x-axis). Robustness calculated across four concentrations for each condition and each strain is shown above the bar plots.*

3.2.4 Robustness quantification using available datasets

The quantification formula (Equation 3) can be used to measure robustness of strains cultivated under diverse conditions, exploiting publicly accessible datasets (19,136,137). In this **thesis**, Equation 3 was applied to evaluate the robustness of five distinct phenotypes across 24 *S. cerevisiae* strains and 29 conditions (**section 2.4:** Dataset 1). However,

acquiring data, particularly through high-throughput methods, is resource-intensive and time-consuming, even if numerous cultivations can be carried out in parallel. Numerous datasets containing diverse phenotypic information are available publicly, enabling the identification of robust strains within specified perturbation spaces. Additionally, robustness quantification applied to yeast deletion libraries can identify specific gene deletions that confer heightened robustness, as demonstrated in **Paper IV** (113).

To prove the flexibility and potential to identify strains with robust phenotypes, Equation 3 was employed on a published dataset encompassing 1011 *S. cerevisiae* strains (**section 2.4:** Dataset 4) (56). The dataset comprised 971 strains, which were grown on agar plates and analyzed in terms of produced biomass across 36 conditions (Figure 3.5) (56). These strains were sourced from 369 distinct locations, representing 24 ecological origins and 312 geographical origins. The top three most robust strains had been isolated from wine amphoras in Georgia. Strains isolated from fermentation environments and palm wine displayed the highest mean robustness; whereas laboratory strains and bioethanol strains exhibited the lowest (Figure 3.5). The tested conditions in this dataset encompassed temperature, sugars, ethanol, salts, metals, and antifungal agents, thereby covering a wide range of perturbations.

The relationship between robustness and ploidy was also investigated with the same dataset (Figure 3.5, right part). Diploid strains exhibited the highest mean robustness, although at higher ploidy values, R values showed less variation and lower means (potentially due to fewer data points available for those categories). Application of Equation 3 to the dataset described above, with 1011 strains (**section 2.4:** Dataset 4) revealed that higher ploidy was not necessarily associated with higher robustness, as observed also when plotting robustness data from the evolution experiment in **Paper V**. There, the haploid parental strain, S288C, exhibited higher robustness than the Ethanol Red diploid strain when grown across 20 different media. Given that robustness is associated with genetic and metabolic redundancy, having more copies of certain chromosomes should imply higher robustness (**Review Paper I**). Nevertheless, even if redundancy implied higher robustness, regulatory networks and gene expression patterns would play a role in the robustness of phenotypes in haploid vs polyploid strains.

Utilizing Equation 3 with available datasets enables the identification of strains highly robust across various origins and facilitates the correlation between strain property (for example ploidy) and robustness. Comparable investigations can be conducted using different datasets, such as those containing information on single-cell phenotypes. These studies serve to evaluate population heterogeneity or screen for potential genes associated with robust characteristics (**Paper IV** and **section 3.2**: Table 3.1).

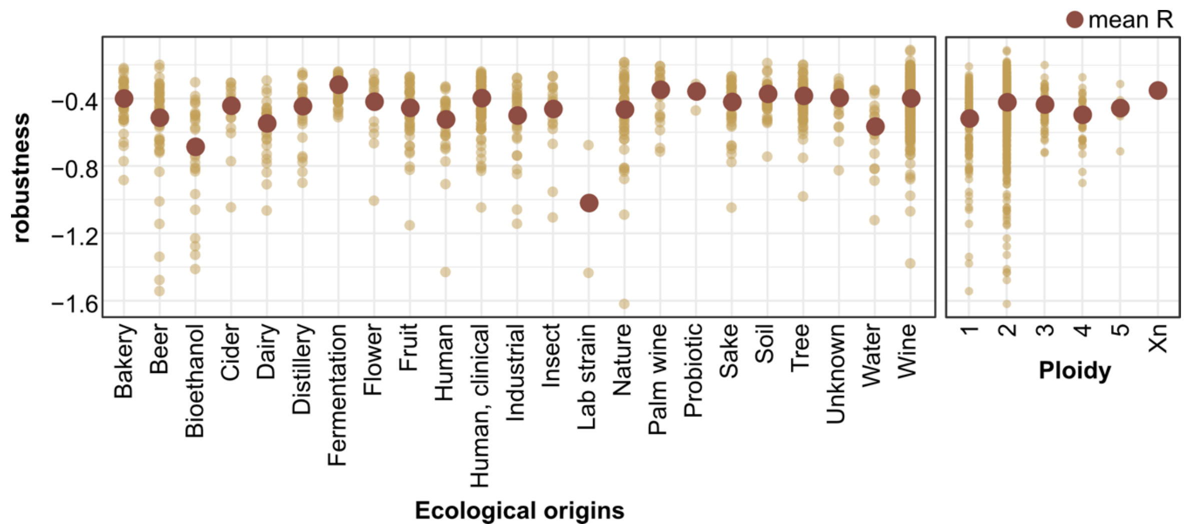


__Figure3.5: Robustness of produced biomass for 1011 S. cerevisiae strains__. Robustness of each of the 1011 strains from Dataset 4 was computed using Equation 3 across 36 perturbations and plotted against various ecological origins (x-axis, left graph) or ploidy (x-axis, right graph). The mean robustness for each ecological origin or ploidy is shown by the dark red dot.

3.2.5 Combined effects of perturbations on robustness

Perturbations are indispensable when measuring robustness, as demonstrated by their influence on the same strains (**Paper III:** Figure 1; **Paper IV:** Figure 6) (7,128). The evaluation of robustness is contingent upon the number and nature of perturbations tested, underscoring the importance of conducting a thorough investigation of the perturbation space before undertaking robustness calculations. Even when many perturbations are tested, including those with low probability of occurrence, removal or addition of pertinent perturbations in a specific process step or environment offers a versatile tool that can be fit to different scenarios. Furthermore, the analysis of a wide range of perturbations could be indicative of a more general robustness mechanism that is not necessarily tied to a perturbation space. A discussion on which perturbations to incorporate into the perturbation space has been addressed in **Paper III** (Figure S2).

Robustness mechanisms are linked to perturbations in three distinct ways:

i) they evolved from previous exposure to similar stressors

ii) they arose from exposure to different stressors

iii) they exist without having any direct association with the perturbation (8).

In the third case, there are only minimal differences in robustness among perturbation spaces, as shown for CEN.PK113-7D met28 in **Paper IV**. Moreover, as discussed in **section 2.2**, both the frequency and magnitude of perturbations have an impact not only on

performance but also on robustness. While laboratory environments often simplify complex real-world conditions by breaking them down into single relevant perturbations, it is important to acknowledge that combined effects between compounds can significantly influence microbial performance (Figure 2.10).

To illustrate the effect of combined perturbations on robustness, the same dataset in which CEN.PK113-7D and Ethanol Red were grown in different combinations of two or three inhibitors and the maximum specific growth rate was calculated (**section 2.4:** Dataset 2; Figure 2.10), was employed to calculate robustness using Equation 3 (unpublished data). Combined effects were simulated by substituting the two individual perturbations with their synergistic action. When examining the combination of NaCl and formic acid, the two single perturbations were excluded from the perturbation space and were replaced by the performance measured in medium containing both stressors. Robustness of the specific growth rate was then quantified using Equation 3 (Figure 3.6).

The trends observed for robustness of the maximum specific growth rate mirrored those seen for overall performance (Figure 2.10), with a noticeable drop in robustness when all three inhibitors were present simultaneously, in particular when combining ethanol with acids, aldehydes, or NaCl. These findings suggest that a high-throughput setup, which tests only single perturbations, may need to be supplemented with combinations of different perturbations to provide a more realistic depiction of robustness in complex environments.

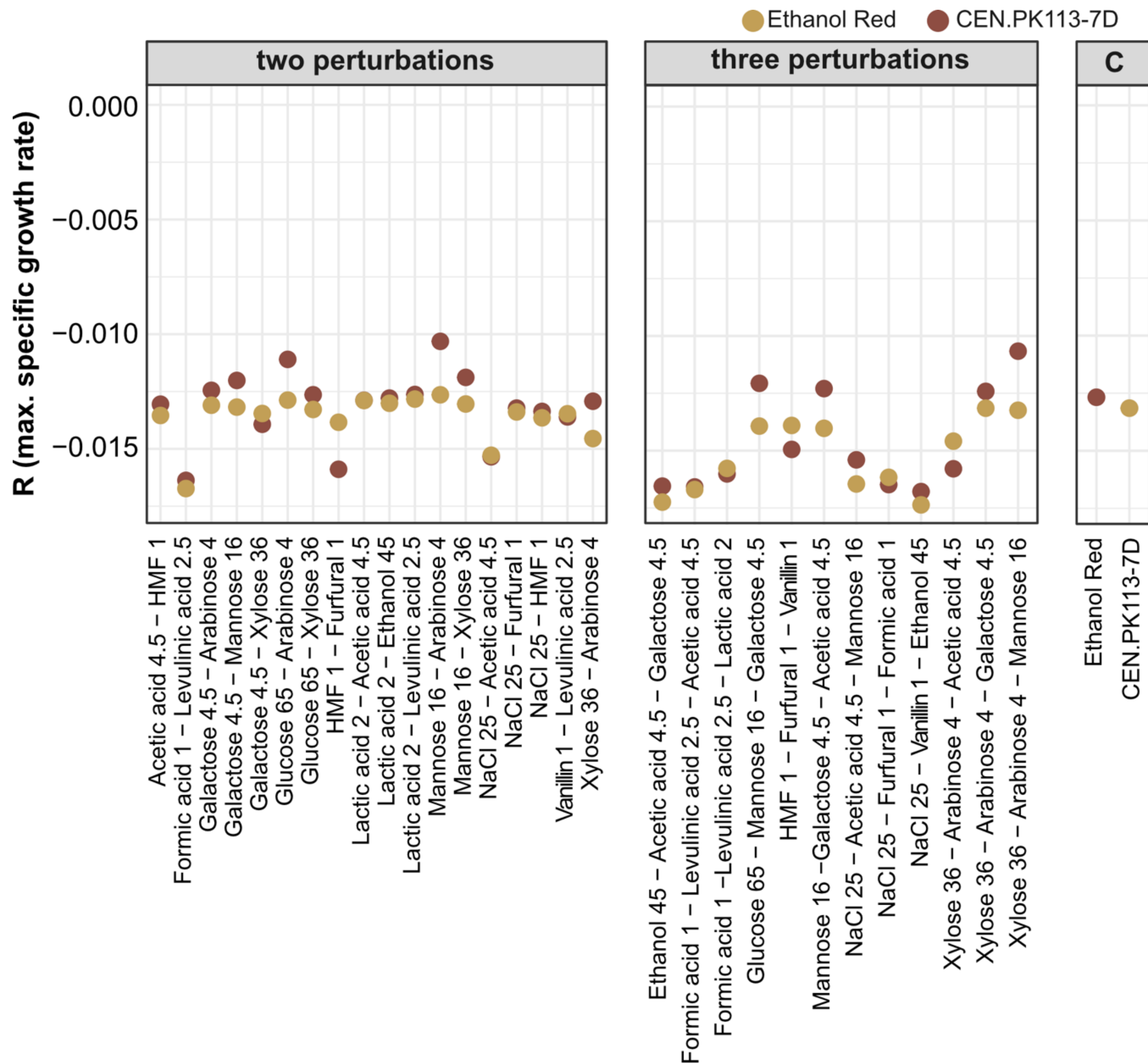


Figure 3.6: ***Effect of combined perturbations on robustness.*** *Robustness of the specific growth rate was computed using Equation 3 and is plotted on the y-axis. Two strains, CEN.PK113-7D and Ethanol Red, are represented by different colors. Robustness was calculated across combinations of two or three perturbations, as plotted on the x-axis. In the control strains, robustness was calculated only for individual perturbations, as shown in the right "C" panel.*

3.3 Robustness vs performance

Equation 3 was designed to quantify robustness according to the definition given in **Review Paper I** and reported in **section 3.1.1**. However, its formulation (**Paper II**) entails that strain performance is not depicted in the formula (high robustness can result both from strains with high or low performance). For instance, biomass data from Peter et al. (2018), who cultivated 1011 yeast strains isolated from different locations and in 36 distinct conditions on agar plates (56) (**section 2.4:** Dataset 4), were used to calculate robustness with Equation 3. When plotting robustness against performance, the data revealed that most

strains exhibited an average mean performance clustered around a robustness value of -0.4 (Figure 3.7). These plots offer a clearer view of which strains exhibit the highest robustness, performance or a good balance between the two (see **section 3.3.2** for detailed optimization analyses). For example, strains CAQ and CAB, isolated from wine preserved in amphoras in Georgia, display the highest robustness but an average mean biomass around 0.3 (values normalized to growth on YPD). Conversely, strains CQG, BMK, BMH, CQH, and BMF, isolated from cocoa bean fermentation in West Africa and clinical samples in Italy, exhibit some of the highest performance and robustness. Finally, SACE_GAP and ACQ isolated from wine in Chile and Russia exhibited the highest performance. These plots can also be used to investigate potential trade-offs between properties.

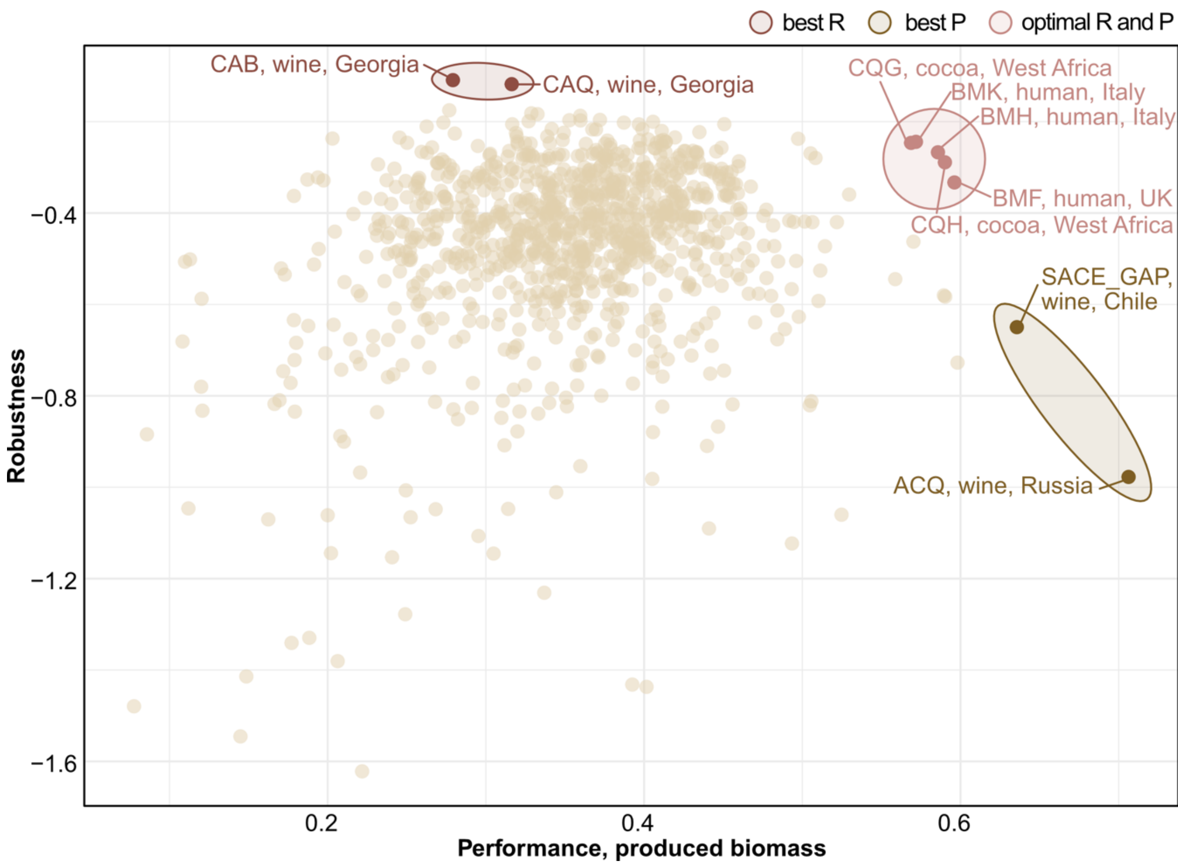


Figure 3.7: Robustness vs performance for 1011 S. cerevisiae isolates. *A total of 1011 yeast strains were cultivated across 36 conditions, and the produced biomass was calculated. Robustness was computed across the 36 conditions for each strain using Equation 3 and plotted on the y-axis against the mean performance of each strain (x-axis). Three shaded areas identify strains with high robustness, performance or with the best compromise between robustness and performance. The code identifying each strain, together with the ecological and geographical origin, is provided (56).*

3.3.1 Trade-offs

By plotting robustness against performance (Figure 3.7), it becomes apparent how most strains engage in a trade-off between high performance and robustness of the same phenotypes. A system cannot maximize all objectives (e.g., specific growth rate, product yield, biomass, and specific growth rate across many conditions) simultaneously due to both biological and thermodynamic constraints (138,139). Trade-offs can arise among performance values of different phenotypes as well as between performance and robustness of the same phenotype (**Paper III**). In large-scale production settings, microorganisms often exhibit a trade-off between production and growth. Typically, after a certain number of generations (often around 100), high-producing cells are outcompeted by non-producing ones. The latter generally display higher specific growth rates, probably due to mutations that have been selected for phenotypes other than product yields (140).

3

Past studies have revealed a trade-off across 62 fungal species between the rate of hyphal growth and the production of melanin, a compound in the cell wall of fungi, which provides resistance to UV light, desiccation, pathogens, and osmotic stress (141). Trade-offs have been observed in bacteria between lag phase and specific growth rate in fluctuating environments (142). Trade-offs between the rate of ATP production and its yield in heterotrophic organisms have revealed even more intricate mechanisms such as the transition from unicellular to multicellular organisms following a rise in ATP yield during respiratory metabolism (143). Moreover, a trade-off exists between the maximum specific growth rate per individual *r* and the carrying capacity (biomass yield) *K*, and is known as the *r/K* selection theory (144).

The trade-offs listed in the previous examples have been supported by phenotypic data collected in this **thesis**; they include specific growth rate, lag phase or cell-dry weight (**Paper III:** Figure S1). The trade-offs observed using data from **Paper III,** with 24 strains cultivated in 29 conditions (**section 2.4:** Dataset 1) revealed a positive correlation between end of cultivation cell dry weight and the maximum specific growth rate. Interestingly, this pattern, which is the opposite of a trade-off, became more pronounced under stressful conditions, such as exposure to ethanol and aldehydes (**Paper III:** Figure S1). This confirms previous observations, where positive correlations were detected in stressful environments, while trade-offs were seen in nutrient-rich media such as YPD (144).

The dataset generated in **Paper V**, was used to further investigate trade-offs (**section 2.4:** Dataset 6). The dataset contained information on generation time (transformed in maximum specific growth rate) and produced biomass of the parental strains Ethanol Red, CEN.PK113-7D, and S288C cultivated in 20 different media on agar plates. Spearman correlation coefficients between the two phenotypes for each group of conditions, were calculated (Figure 3.8). In line with what observed above with Dataset1, negative correlations were only observed in the case of Ethanol Red cultivated with HMF and

CEN.PK113-7D cultivated without pH buffering or trace metals. Therefore, positive correlations in this case were observed not only in the case of stressful environments but also in control conditions such as Delft with 2% glucose.

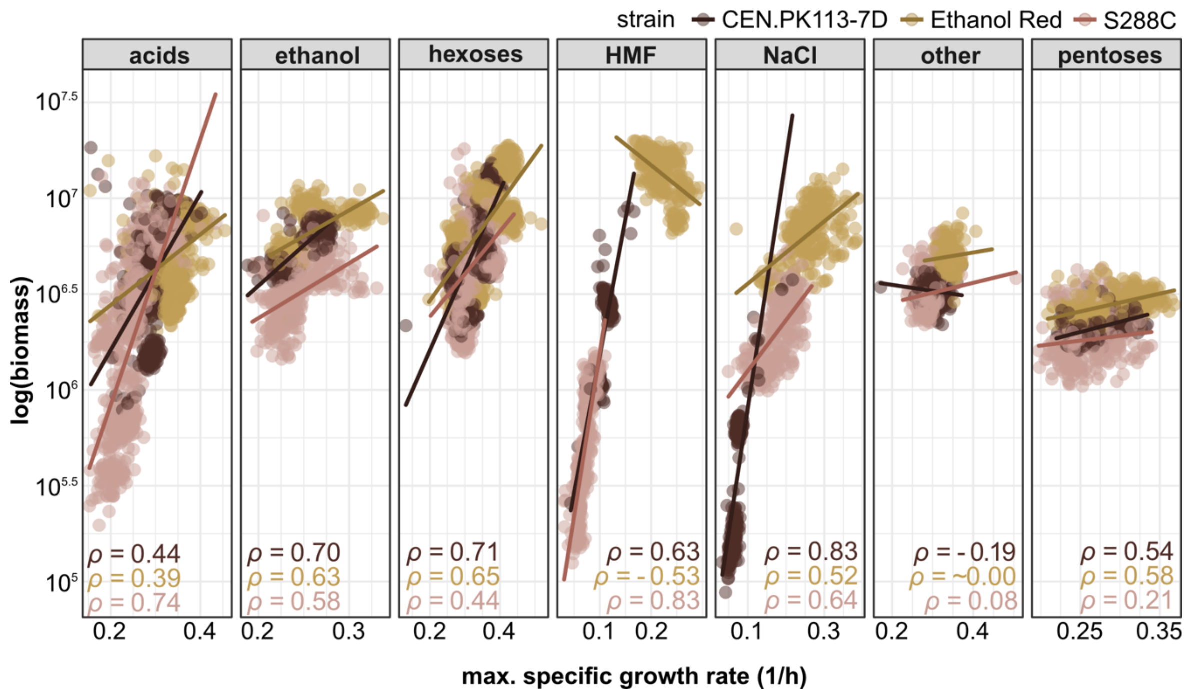


Figure 3.8: Correlation between maximum specific growth rate and produced biomass in three S. cerevisiae strains*. The produced biomass was plotted on a logarithmic scale (y-axis) against the maximum specific growth rate (x-axis) for three S. cerevisiae strains (colors) and seven groups of conditions (columns). Each facet displays the Spearman correlation coefficient rho "ρ". The significance was not represented but the p-value was ≤ 0.0001 for all correlations except for Ethanol Red and S288C in the "other" group, whereby the p-value was not significant. A linear regression line denotes the direction and strength of the correlation. The data are extracted from the parental strains phenotypic assays (**Paper V**).*

Trade-offs were observed with respect to performance of different phenotypes but also between robustness and performance of the same phenotype, as discussed in **Paper III**. However, contrary to expectations based on existing literature, data from various phenotypes revealed that trade-offs between performance and robustness were not always present. In some cases, such as the specific growth rate, the correlation between performance and robustness was positive, likely due to evolutionary mechanisms that have optimized both properties (**Paper III**). While correlation studies can provide valuable insights when assessing large datasets with phenotypic data, it is important to remember that these studies are limited to the experiments performed and should not be relied upon solely as evidence of trade-offs or to confirm causal mechanisms.

3.3.2 Pareto fronts

A useful tool to investigate trade-offs and multi-optimization strategies among different quantities is Pareto optimality. The Pareto front is a geometrical shape that represents Pareto optimality and has often the shape of a curve. It can be used in multi-objective optimization studies to identify the best possible compromise between conflicting objectives, such as performance and robustness (145–147). Based on Figure 3.7, in which the performance and robustness of produced biomass were plotted for 1011 *S. cerevisiae* strains (56) (**section 2.4:** Dataset 4), a Pareto front was plotted by delineating the dominance region of the plot, i.e., where one variable dominated or had greater influence over another variable (Figure 3.9). The Pareto front links two archetypes, representing the maximum of each property: one strain with maximum performance (ACQ: isolated from Russian wine) and another with maximum robustness (CAB: isolated from Georgian wine), via a line connecting all other strains that exhibit varying degrees of optimization between the two properties. In this context, the same results indicated by circles in Figure 3.7, were now emphasized by the Pareto Front.

Pareto optimization strategies applied to genome-scale metabolic models have been used before to propose gene knockouts that would improve both specific growth rates and product biosynthesis (148). Pareto front analysis has been instrumental in understanding why certain *E. coli* strains were not optimized and, therefore, did not appear on the Pareto front. The strains prioritized robustness in environments subjected to perturbations, optimizing for stability over peak performance. Specifically, these cells adjusted their metabolic flux to sustain growth across various conditions, which resulted in a compromise, whereby they were not as close to the Pareto front (149). Robustness has not been yet integrated in the Pareto optimization analysis. Now, typical optimization studies are performed between microbial production and growth, for example. Coupling Pareto optimization with conditional robustness algorithms (technique that quantifies how perturbations affects cellular parameters and metabolic outputs of biological networks) could optimize robustness objectives together with performance in different perturbation spaces.

The phenotypic data (produced biomass vs maximum specific growth rate) of 1044 strains from **Paper V** grown in 20 conditions on agar plates (**section 2.4:** Dataset 6) was exploited to show the application of the Pareto front in dynamic settings. In this case, evolved *S. cerevisiae* strains displayed different Pareto fronts for the traits produced biomass and generation time. The shift in the Pareto front showed that cells had adapted over time, not only in terms of produced biomass or specific growth rate, but certain populations had successfully optimized both variables simultaneously.

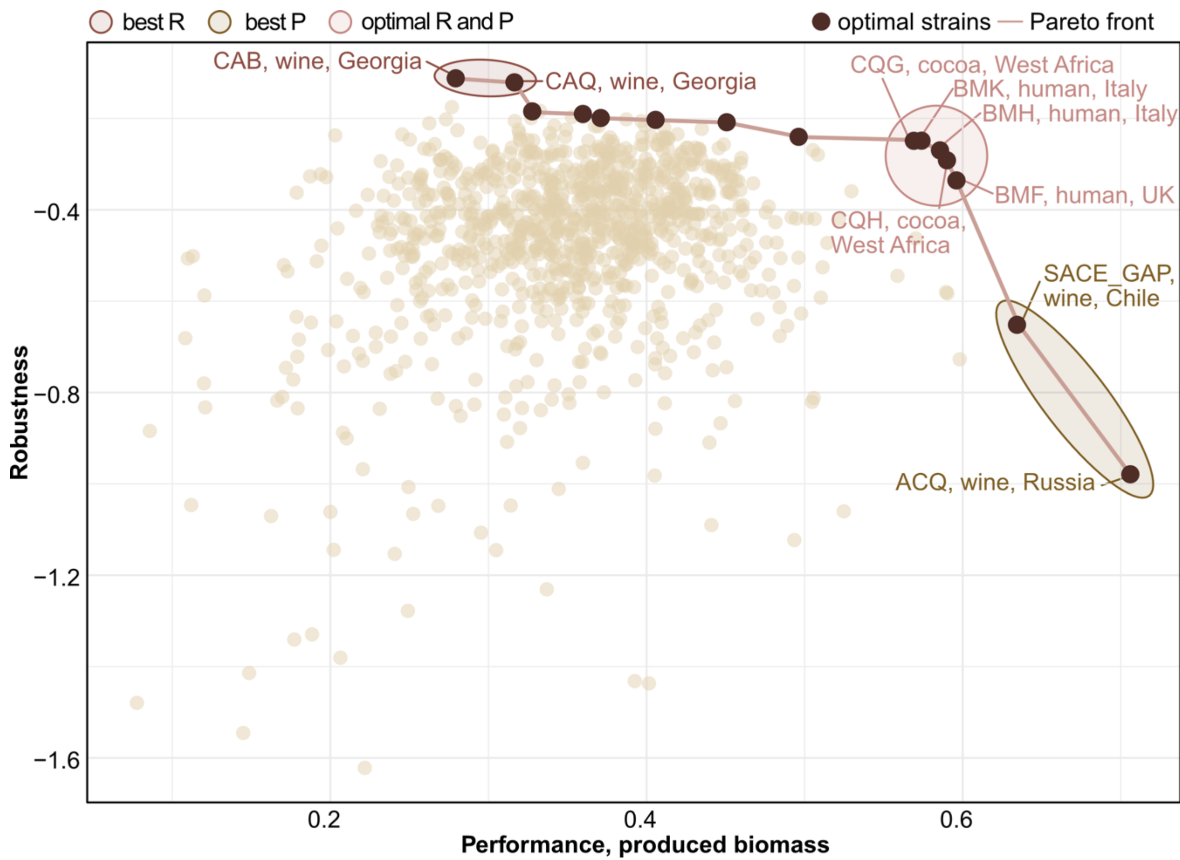


Figure 3.9: Pareto front for robustness and performance of produced biomass by S. cerevisiae strains. *A total of 1011 yeast strains were cultivated across 36 conditions, and the produced biomass was calculated relative to the produced biomass on YPD agar. Robustness was computed across the 36 conditions for each strain using Equation 3 and plotted on the y-axis against the mean performance of each strain (x-axis). The best robustness or performance and the best compromise between the two were previously identified by the shaded area in Figure 3.7. The Pareto front is highlighted by a line and the best, optimal strains in terms of robustness and performance corresponds to the points along this line. The code identifying each strain, together with the ecological and geographical origin, is provided (56).*

In Figure 3.10 the Pareto front was built using only the parental strains (left) and then only the evolved strains (right). A comparison of the plots from parental or evolved strains, highlighted the shift in the Pareto Front. For Ethanol Red, the Pareto front was almost identical in the two cases, indicating that the strain did not further optimize its performance (as concluded in **Paper V**). By confronting the Pareto front of different strains both before and after evolution we might be able to gain insights on the strain adaptability, as reported previously in *S. cerevisiae* (150). Furthermore, by including a high variety of strains (e.g., Dataset 4 (56)), showcasing natural diversity in the Pareto front analysis, it might be possible to delineate realistic biological limits to optimization, which could direct strain design towards tolerance and even simultaneous optimization of different traits.

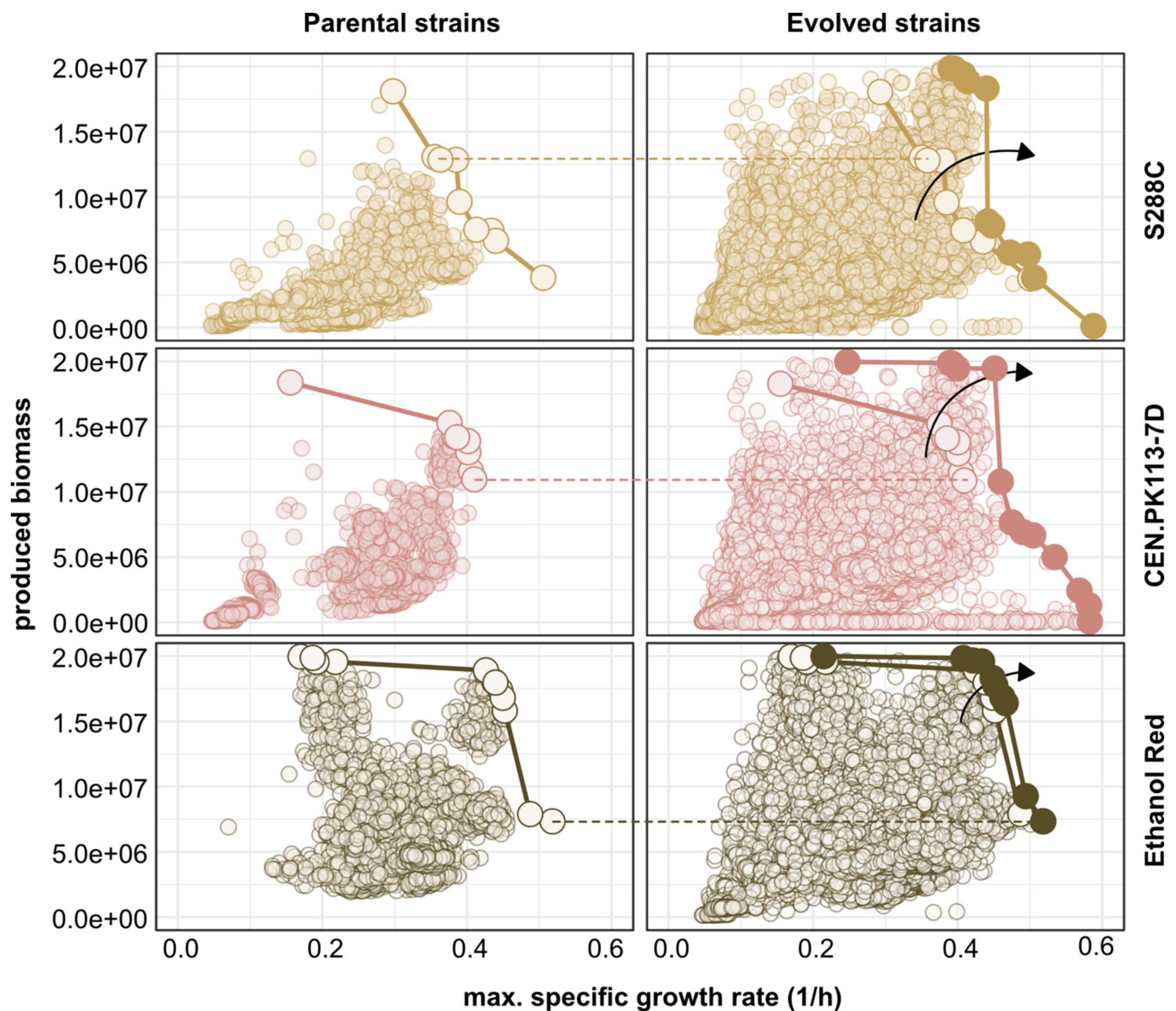


*Figure 3.10: **Dynamics of Pareto fronts with parental and evolved S. cerevisiae strains.** The maximum specific growth rate (x-axis) is plotted against the produced biomass (y-axis) for three S. cerevisiae strains (each row). The panels to the left show the parental strains' fitness; those to the right show the evolved strains' fitness. The Pareto front for both parental and evolved strains is represented by a continuous line joining together large dots. The dotted horizontal line for each strain indicates where the Pareto front of the parental strain was superimposed on the evolved strains' graph, and the black arrow points to the shift in the front after evolution.*

Main Points from Chapter 3

- The Fano factor-based quantification is preferred to others as it is frequency independent, dimensionless, and free from arbitrary control conditions.
- The quantification formula can be used for many applications, for example to study trade-offs or to identify strains with robust phenotypes for industrial purposes.
- Pareto fronts are a useful tool to identify strains that show the best compromise between two traits or objectives, such as performance and robustness.

Chapter 4. Microbial robustness concepts applied to evolution and genomics

Chapter 4 presents a summary of how genetic markers linked to microbial robustness can be identified. Genetic markers were derived from existing literature and by applying Equation 3 to different screens of yeast deletion collections. The latter part of the chapter discusses the evolution of microbial robustness, and employs fitness landscapes as a theoretical framework to elucidate evolutionary outcomes. Additionally, the chapter touches on history-dependent behaviors in relation to the results presented in **Paper III**.

The development of a methodology to quantify microbial robustness has enabled its quantification across different strains and perturbation spaces. While previous chapters primarily explored robustness in terms of measurable phenotypes such as specific growth rates or produced biomass during cultivation; this chapter focuses on the intracellular aspects of robustness. Equation 3 was applied to yeast deletion libraries containing fitness data to explore metabolic hot spots connected to robustness in different perturbation spaces (**Paper IV**). Additionally, it allowed the investigation of how different evolution set-ups influenced the robustness of produced biomass by three *S. cerevisiae* strains (**Paper V**).

4.1. Genetic markers of microbial robustness

The stability of microbial phenotypes across diverse environmental settings (i.e., robustness) is achieved via a collaborative action of core intracellular mechanisms, spanning from complex metabolic pathways and their regulation to gene expression and transcription-related events (8). Microbial robustness is characterized by three principles: i) redundancy, ii) modularity, and iii) control strategies (as discussed in **Review Paper 1:** Figure 1). Redundancy refers to different genes and metabolic pathways covering the same function (151). It can also refer to the complex wiring of different transcription factors. Transcription factors are proteins that bind to a specific DNA sequence and regulate the rate at which genetic information is transcribed from DNA to mRNA. For example, single transcription factors may bind to the same regulatory DNA element or, conversely, transcription factors from different families may interact with a single regulatory element (152). Modularity refers to the modular organization of biological networks (153). Control strategies, such as feedback loops, detect variation in a specific output and apply a corrective action in the input, which immediately compensates for the sensed variation (154). Examples of feedback loops adopted as a control strategy are the lactose operon in *E. coli* or galactose utilization in *S. cerevisiae* (155). These characteristics act together, making it difficult to pinpoint specific pathways or genes that could be uniquely responsible for robustness.

4.1.1. Previously investigated markers

Previous studies have reported a variety of genes, which act as “robustness factors” or “phenotypic capacitors” to ensure low phenotypic variability across perturbations. One of them is heat shock protein 90 (Hsp90), a molecular chaperone and key regulator of proteostasis under both standard and stressful conditions. Chaperones interact with other proteins to make them functionally active. Hsp90 is involved in protein folding, binding of ligands to their receptors, and assembly of multiprotein complexes (156,157). Studies on Hsp90 from *Arabidopsis* and *Drosophila* have shown that its disruption correlates with higher phenotypic variation (158). However, due to lack of comparison with other genes and poor understanding of Hsp90 buffering mechanisms, further investigation of Hsp90 and its ability to lower phenotypic variance is needed (7).

More recent studies have used high-throughput approaches to assess morphological changes in single-gene deletion strains and discover phenotypic capacitors (159). Siegal et al. identified phenotypic capacitors enriched in the following Gene Ontology terms: chromosome organization, DNA integrity, RNA elongation, and response to stress. The identified capacitors include CCR4, whose deletion caused irregular colony morphology (160), SWI6, whose deletion caused variability in cell size in liquid media (161), and FUS3, whose knockout shows cell-to-cell variation in response to pheromones(162). Furthermore, the same study revealed that phenotypic capacitors acted as hubs (i.e., highly connected nodes) in protein-protein interaction networks. Hsp90 and its homologue Hsp70-SSE1 are both phenotypic capacitors and network hubs. It has also been shown that deletions or mutations of network hubs relate with specific growth rate variation in *E. coli* in response to different environmental perturbations (163).

To dive into the discovery of robustness markers, I applied Equation 3 on publicly available yeast deletion collection screens to quantify their robustness.

4.1.2 Yeast deletion collections: a tool for identifying robustness markers

The findings from this **thesis** and previous research suggest there are no universal regulators of microbial robustness (164). However, applying Equation 3 to single-deletions libraries could reveal unknown genetic or metabolic determinants of robustness. In this regard, datasets from yeast knockout (YKO) libraries, as well as other phenotypic datasets, have proven valuable in elucidating genotype-to-phenotype relationships, phenotypes correlations, and transcriptional responses to various stressors (139,165,166). Hence, perturbation experiments, in which different strains or YKO collections are exposed to multiple perturbations, remain a primary tool for exploring the mechanistic basis of robustness.

A dataset containing more than 4000 single-gene deletions in *S. cerevisiae* (derived from strains BY4741 and BY7092) was grown under 14 conditions to investigate genes and mechanisms correlated with robustness (113)((**Paper IV**) (**section 2.4:** Dataset 3). The robustness quantification formula (Equation 3) was applied to the dataset and fourteen gene deletions with the highest or lowest values of either robustness or fitness were selected and replicated in *S. cerevisiae* CEN.PK113-7D. The fourteen deletion mutants, the industrial strain Ethanol Red, and the parental strain CEN.PK113-7D were grown in three distinct perturbation spaces each with 16 perturbations (**Paper IV**). The three perturbation spaces mimicked beer fermentation (beer perturbation space), 2G bioethanol production (lignocellulose hydrolysate perturbation space), and the original conditions containing antifungal agents and sugars (Costanzo perturbation space) (113). This strategy revealed gene deletions that resulted in robust phenotypes in one perturbation space, but not in others (**section 2.4:** Dataset 5).

Deletion of the SMA2 gene (**Paper IV:** Figure 6) resulted in a robust specific growth rate in the beer and Costanzo perturbation space, but not in the lignocellulose hydrolysate perturbation space. In contrast, deletion of MET28 (**Paper IV:** Figure 6) resulted in a robust but overall lower specific growth rate in all three perturbation spaces. MET28 participates in sulfur metabolism, influencing glutathione biosynthesis and other metabolic processes such as DNA replication (167,168). Glutathione protects against oxidative stress during lignocellulose ethanol production (169,170). Deleting MET28 destabilizes the Met4 complex and sulfur metabolism in *S. cerevisiae*. At present, it remains unclear whether MET28 alone contributes to robustness, because genes related to it, including CBF1 or MET4, were not examined. Nevertheless, the findings in **Paper IV** suggest that sulfur metabolism may be responsible for robustness mechanisms.

Akin to the approach taken in **Paper IV**, datasets featuring phenotypic data of YKO strains grown under different experimental conditions can be used for robustness analysis. Turco and colleagues published a collection of more than 14,500 YKO screens describing 6731 phenotypes and 7536 experimental conditions (171) and referred to it as the Yeast Phenome. In one of these screenings, a subset of the YKO collection encompassing 604 single-deletion strains (114) (**section 2.4:** Dataset 7) was grown on synthetic complete medium agar plates supplemented with 3 mM arsenite ([As III]; $NaAsO_2$), 4 mM arsenite, 0.25 ug/mL rapamycin, 400 ug/mL paraquat (methylviologen; N, N-dimethyl-4-4′-bipiridinium dichloride), 50 uM cadmium chloride or 1.25 M NaCl. Growth was monitored using the scan-o-matic set-up (**section 2.4:** high-throughput methodologies) and the produced biomass and doubling time were calculated from the resulting growth curves (172).

When Equation 3 was applied to calculate robustness of the produced biomass and generation time across the tested conditions (**section 2.4**: Dataset 7), some strains carrying

single deletions displayed notably lower robustness than the parental strain (Figure 4.1.a). One of them was *oca1*, which encodes a tyrosine phosphatase required for cell cycle arrest in response to oxidative damage of DNA, and exhibited 86% lower robustness for generation time than the wild-type strain. Two other candidates were *met3*, which corresponds to an ATP sulfurylase, and *rpS21a*, which encodes a component of the small (40S) ribosomal subunit (173–175).

Robustness of produced biomass (from analysis of robustness using Dataset 7) revealed two more targets for robustness: *rpL27a* and *ssa1*, which encode a protein of the large ribosomal subunit and a member of the Hsp70 family, respectively (176,177). Although the mutants highlighted by robustness analysis of this dataset (**section 2.4:** Dataset 7) differed from those studied in **Paper IV**, significant functional and regulatory interactions were found among the translated proteins (e.g., Rps14a from **Paper IV** interacts with Rps21a from Dataset 7 (114). The interactions were confirmed by the analysis conducted with the STRING database. STRING systematically compiles genomic associations and protein interactions, assigning confidence scores to assess their significance, thus facilitating the understanding of protein networks and functional relationships (178) (Figure 4.1.b).

Despite differences in the examined conditions between Dataset 7 and Dataset 5 (**section 2.4**), shared genes associated with robustness were identified. Sulfur metabolism, heat shock proteins, and ribosomal proteins have emerged as the likeliest effectors of robustness. In the above investigation, robustness was evaluated in terms of generation time and produced biomass, with different genes identified for each phenotype. This finding underscores the importance of examining different phenotypes, because various metabolic processes may underlie robustness (as discussed in **Papers II and III**, as well as **section 2.2.2**). Additionally, such analysis could serve as a straightforward illustration of how robustness quantification can be applied to YKO collections.

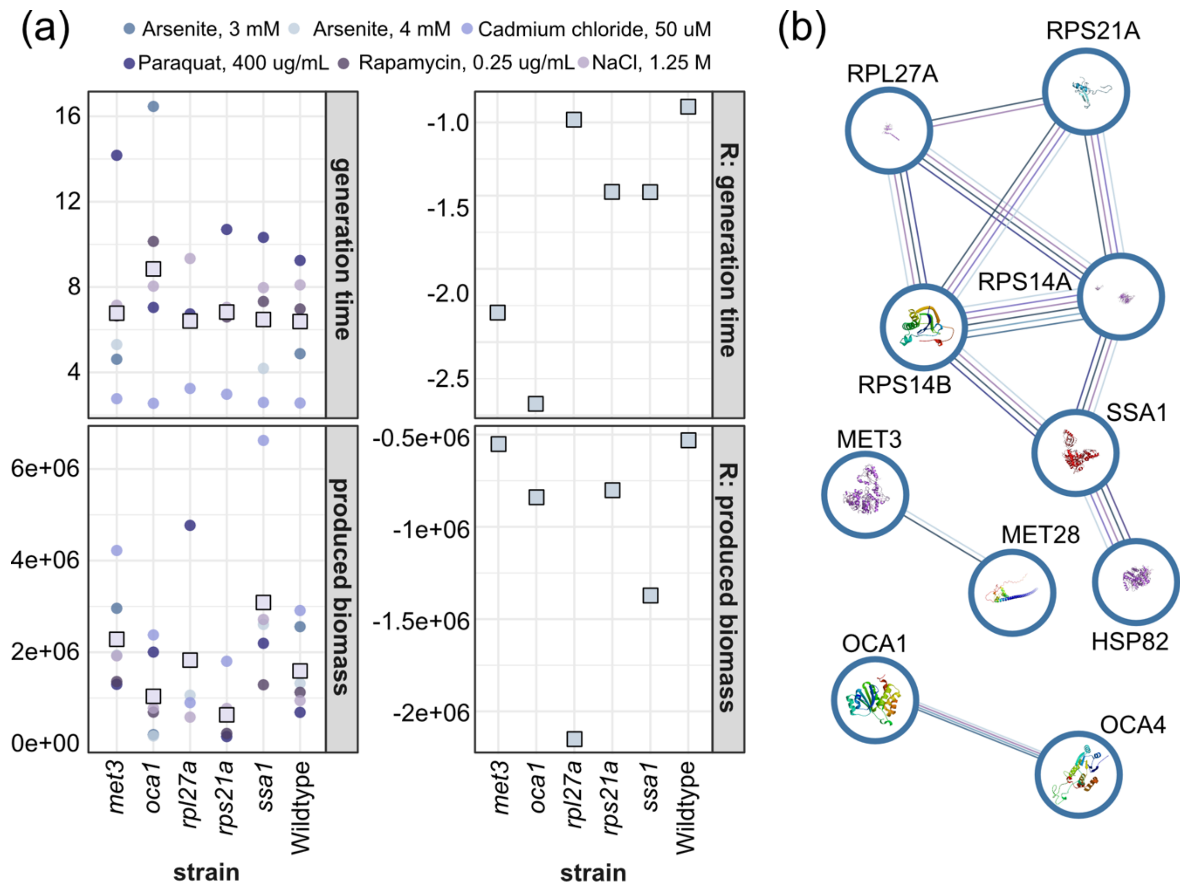


__Figure 4.1: Performance and robustness of single-deletion mutants and protein interactions.__ a) Five mutant strains carrying a gene deletion plus the wild-type (x-axis) were plotted against the performance (left) and robustness (right) of produced biomass and generation time (y-axis). The dots correspond to the mean of replicates (n=468) in each environment. The mean of the six environments is represented by a square. Hues represents the different perturbations. b) STRING network with relevant robustness markers from Dataset 5, Dataset 7, and literature. Each circle represents a gene and its translated protein, while the lines connecting the circles represent different types of interactions (e.g., gene fusion, co-expression).

4.2 Evolution of microbial robustness

Section 4.1.2 showed how the investigation of robustness markers in YKO collections pointed to potential key metabolic mechanisms. However, to gain a broader understanding of robustness, extensive screenings across multiple conditions and phenotypes are necessary. Analysis of more complex mechanisms is limited by the amount of deletions. Consequently, in **Paper V**, the mechanisms underlying robustness were explored from an evolutionary perspective. The hypothesis that served as the foundation for the experimental design originated from the definition of robustness:

If robust strains can consistently perform across various conditions, can fluctuating environments serve as a selective pressure in evolving robustness within a laboratory setting?

To address this hypothesis, three *S. cerevisiae* strains (S288C, CEN.PK113-7D haploid, and Ethanol Red diploid) underwent evolution in both stable and fluctuating environments for 300 generations. Then, the parental strains as well as the evolved samples were assessed in terms of fitness (produced biomass) and the genome of all evolved populations was sequenced (**Paper V**). To align with previous research on evolution, in this **section**, the performance of a phenotype will be referred to as fitness. Before delving into the results of the study, a brief overview of the method is provided.

4.2.1 Adaptive laboratory evolution

Adaptive laboratory evolution (ALE) is commonly employed in biological research to explore evolutionary processes and to improve strain fitness. The fundamental principle of evolution is that natural selection favors the survival and reproductive success of organisms best suited to a given environment (179). As mutations accumulate over time, they alter the fitness of a population. Mutations that negatively impact an organism's survival (deleterious mutations) tend to be eliminated through natural selection; whereas those that have a positive impact (beneficial mutations) are preserved. In reality, even though harmful mutations are typically selected against; in some cases, beneficial mutations might also mask deleterious ones. The emergence and maintenance of mutations is influenced by several factors, including the effect on fitness of new mutations, population size, and ploidy (180,181).

Microorganisms are particularly amenable to laboratory evolution owing to their large population size (i.e., a larger number of individuals provides a greater chance for beneficial mutations to arise and potentially be selected), fast specific growth rates, and ease of cultivation (179). Different set-ups can be used in laboratory evolution experiments. Batch transfers in ALE, whereby cell cultures are propagated in fresh media by serial dilutions, can increase tolerance to specific stressors (18). In ALE experiments using batch set-ups, the growth phase at the time of transfer influences the selected phenotype. Cells are generally transferred in exponential phase to avoid long stationary phase adaptation, which could give rise to mutants that are specialized in utilizing nutrients from dead cells (182). Alternatively, cells can be maintained at constant specific growth rates inside chemostats, where environmental variables are tightly controlled (183). Chemostat evolution with *E. coli* om glycerol-limited environments has been used to improve its growth and biomass yields on glycerol (184). In **Paper V**, batch transfers were used, cells were transferred in late exponential phase or at the beginning of stationary phase. Given the variety of conditions (15 distinct media) and the multiple strains involved, it was difficult to pinpoint the exact

growth phase at transfer. In preliminary experiments, when the growth curve of each strain was monitored under all conditions, the transfer was planned in mid-exponential phase.

Evolutionary experiments can enhance tolerance within a specific medium. Typically, that is done by employing batch transfers with incrementally higher levels of the stressor. Populations selected under these conditions evolve mechanisms that allow them to withstand stressors, while proliferating at elevated specific growth rates. Adaptation to a particular environment can hinder performance in a different setting (185) and has been attributed to antagonistic pleiotropy (186). Trade-offs can be avoided by using fluctuating environments. However, strains evolved under fluctuating conditions tend to exhibit lower fitness across all environments, in contrast to those evolved under stable conditions (187). Nevertheless, a recent study has demonstrated that, on rare occasions, fluctuating selection can result in populations that are more fit to a certain environment than those specialized for it (188). In **Paper V**, fitness data revealed that populations evolved under fluctuating conditions exhibited higher fitness than those cultivated in constant environments.

4.2.2 Phenotypic assays in evolution experiments

The fitness of evolved populations can be assessed using various methods (189). The three main indicators are maximum specific growth rate in different media, minimum inhibitory concentration of stressors, and competition assays. Competition assays are favoured in evolutionary biology because they account for variations throughout the culture cycle, including lag phase and stationary phase dynamics, which are not captured by maximum specific growth rate alone (189). However, these assays are challenging to perform when using multiple replicates and high-throughput settings. In **Paper V**, phenotypic assessments were conducted using a high-throughput platform called scan-o-matic (172) (**section 2.4:** High-throughput methodologies). Briefly, this system enables parallel growth of 100,000 populations on agar plates (**Paper V**: Figure 1), which is its principal advantage over liquid cultures and competition assays. Colony growth is tracked by scanners that record changes in color and size, with spatial normalization to offset nutrient gradients on the plates (172).

During evolution experiments in this **thesis** work, which involved μL-volumes of liquid cultures, phenotypic evaluations were preceded by comparisons of generation times in both solid and liquid media (**section 2.4:** High-throughput methodologies) for randomly chosen samples from the evolutionary plates of three *S. cerevisiae* strains. This test was performed to determine if high-throughput phenotypic systems could be used to evaluate the fitness of evolved samples, as liquid cultivation was not possible due to the high volume of samples and conditions. Five media were tested: YPD, Delft, and Delft supplemented with either acetic acid (0.2%), HMF (0.1%) or NaCl (2.5%). Solid media resulted in shorter and more consistent generation times compared to liquid (Figure 4.2, unpublished data). This can be explained by the cells initially sensing the stressor on the agar surface but, as colonies grow

both outward and upward the concentration of the stressor for the cells inside of the colony is lowered. Consequently, for the phenotypic assessments in **Paper V**, stressor concentrations were higher than those used during evolution. Correlation tests, comparing generation times on solid and liquid media from the phenotypic assays, suggested a significant overall similarity (p-value < 0.001, unpublished data). Only Delft media exhibited a weaker correlation. Given such significant correlations and the high-throughput set-up, the scan-o-matic system was employed for phenotypic assays of evolved populations.

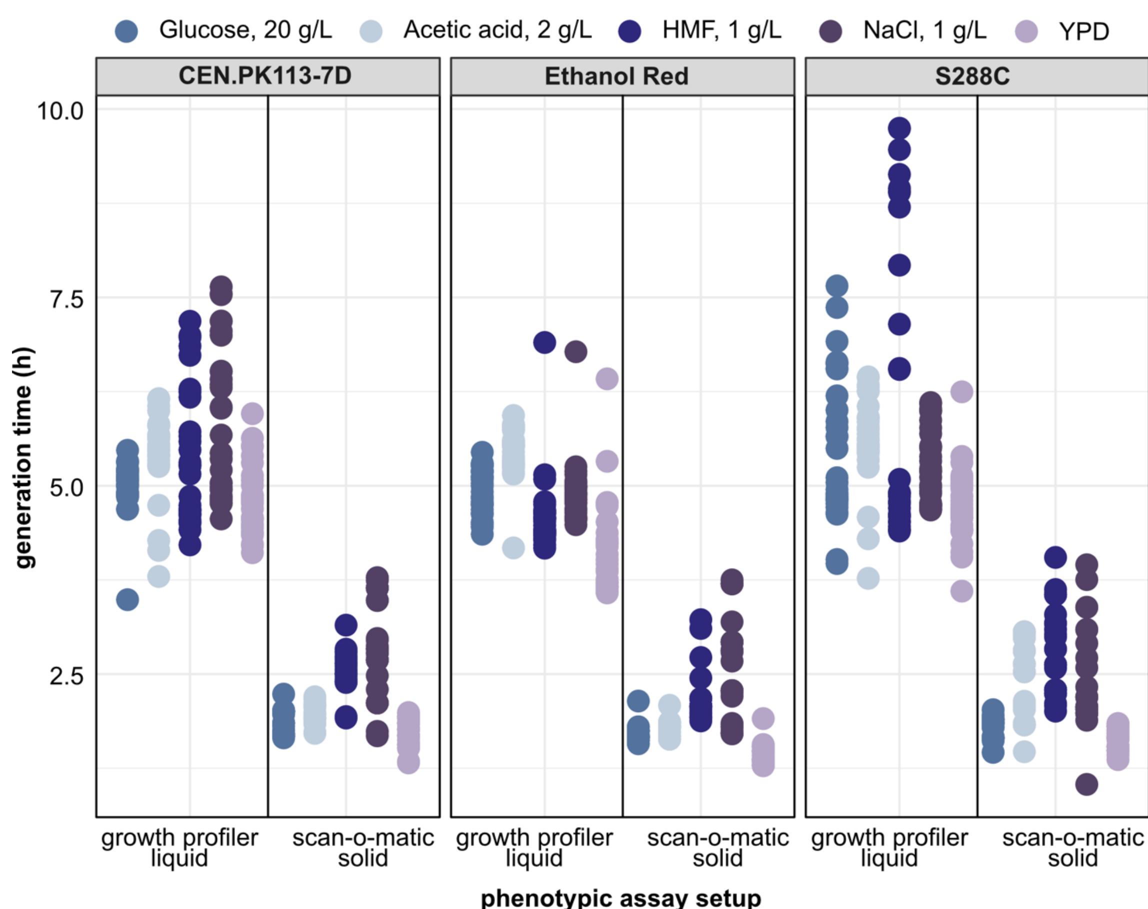


Figure 4.2: Comparison of generation times in solid vs liquid growth. *The generation time was plotted on the y-axis for randomly picked evolved samples (**Paper V**) grown in liquid cultures (growth profiler) and solid medium (scan-o-matic) (x-axis). Each panel represents a different strain. Various colors denote different growth media.*

4.2.3 Fitness and robustness in ALE

In my study, the scan-o-matic system was adopted to evaluate the fitness of evolved populations across 20 diverse conditions, including those encountered during evolution and others with higher concentrations of stressor (**Paper V:** Material and Methods) (**section 2.4:** Dataset 6). Numerous conditions could be tested and a more reliable depiction of robustness was achieved (Figure 4.3). Equation 3 was applied to compare the robustness

of evolved samples and parental strains. Contrary to the hypothesis introduced in **section 4.2**, where robustness would always increase in evolution in fluctuating environments, the results did not uniformly support an increase in robustness following evolution in fluctuating environments. Specifically, while Ethanol Red exhibited higher robustness when evolved under fluctuating conditions (**Paper V**: Figure 4); S288C and CEN.PK113-7D failed to display a significant improvement, despite enhanced fitness after 300 generations (**Paper V:** Figure 3 and Figure 4). The discussion in **Paper V** attributed the observed evolutionary outcomes to strain genetic background, with ploidy likely playing a key role.

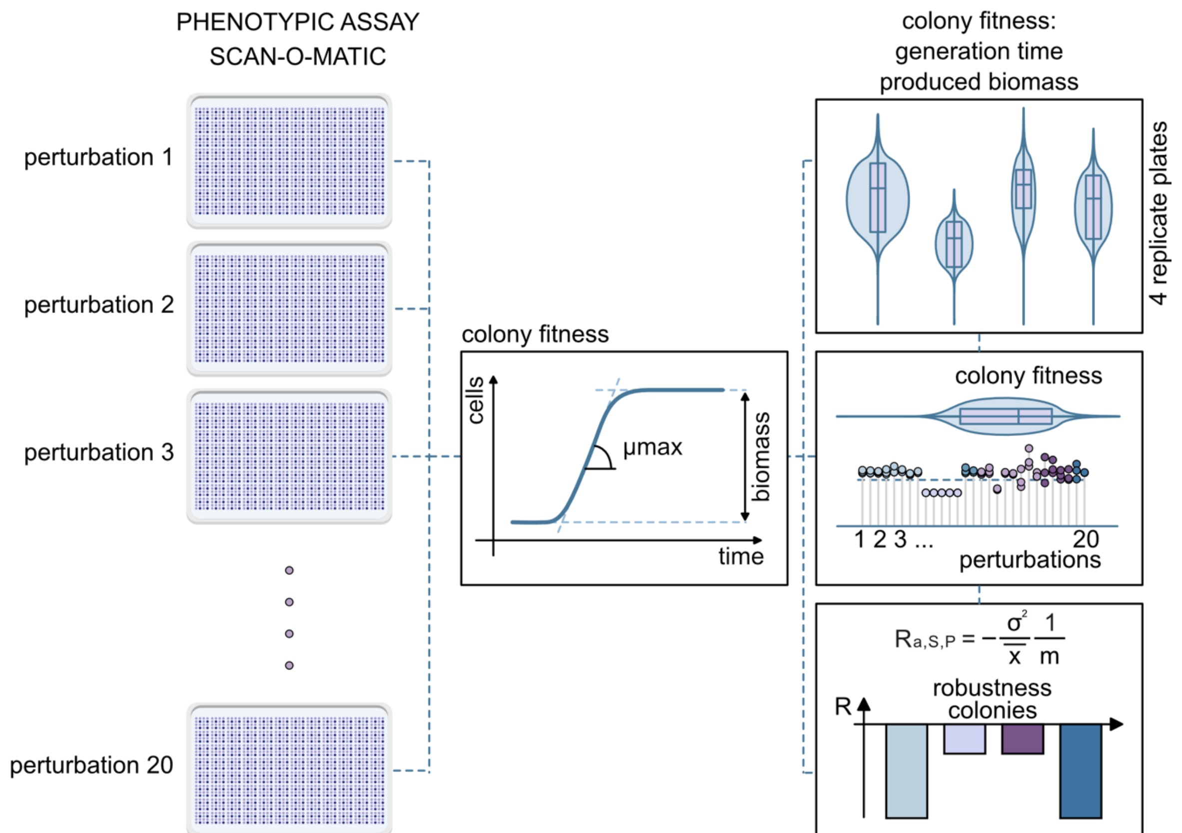


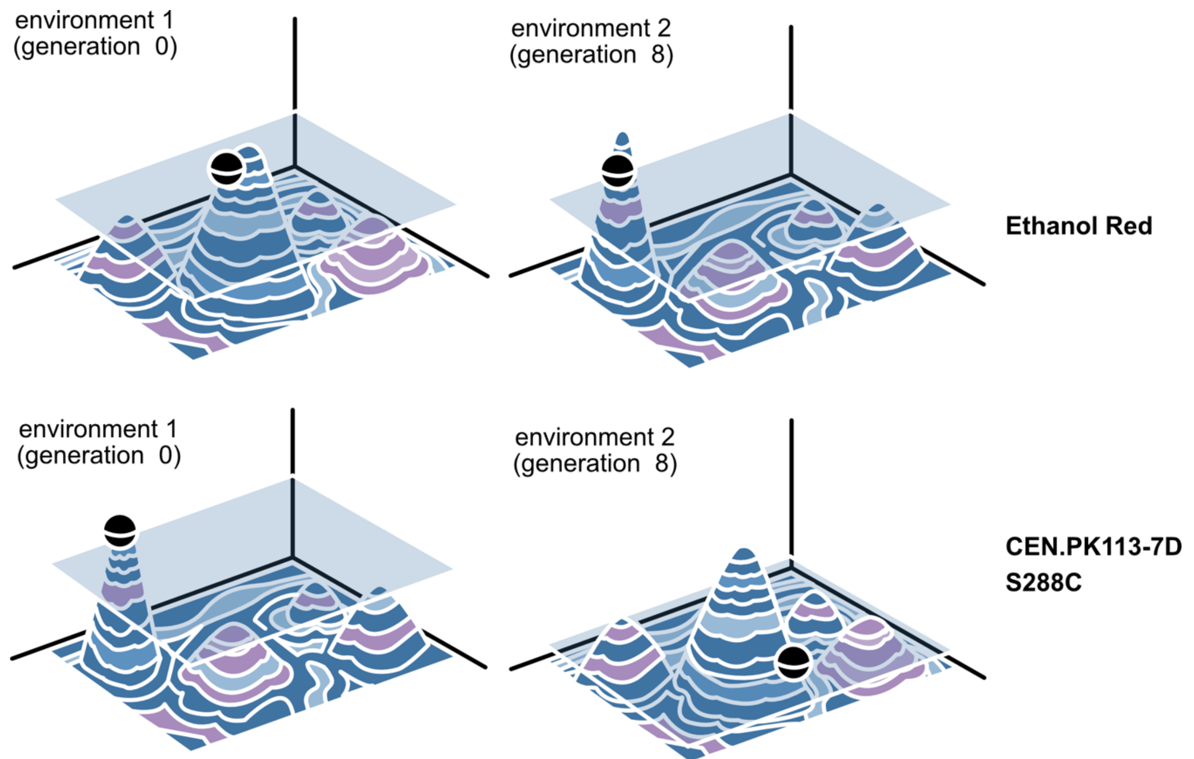
Figure 4.3: Method for calculating robustness from the phenotypic assays of evolved samples (section 2.4: Dataset 6)*. Evolved populations were cultured on agar plates in the scan-o-matic set-up, with four replicates per condition and across 20 different media. Generation time (h) and produced biomass were calculated from the growth curves. Robustness for each population/colony was then quantified using Equation 3, based on fitness data gathered from all replicates.*

Fitness landscapes offer a straightforward method for visualizing the evolutionary trajectory and changes in fitness in relation to a strain's genetic composition. These landscapes are depicted in a three-dimensional plot, with all possible genomic sequence combinations plotted across the x-y plane (genotypes), and fitness represented on the z-axis. In the resulting landscape of peaks (high fitness) and valleys (low fitness), the former correspond to high fitness (190). Landscapes take different forms in different environments.

Every point on these landscapes represents a distinct genotype, with proximate points sharing greater genetic similarity. Mutations drive the organism's evolution through this landscape, typically culminating at a fitness peak during what is termed an “adaptive walk” (191). In landscapes with multiple peaks, a cell population may settle on any peak, usually the one nearest to their evolutionary starting point. In a specific environment, genotypes with initially lower fitness often adapt more rapidly than their fitter counterparts (192). This is attributed to the fact that fit phenotypes sitting at the top of the landscape's peaks are less inclined to descend and traverse another adaptive route; whereas a less fit genotype, possibly nestled in a valley, is more predisposed to adapt and ascend to a peak.

In the evolution experiment of **Paper V**, it is possible that Ethanol Red did not significantly improve fitness because it was already near a fitness peak within each evolutionary environment. Consequently, despite shifts in environments and corresponding changes in fitness landscapes, Ethanol Red may have consistently occupied a position of high fitness (light blue plane, Figure 4.4). During evolution, the strain reached a similar fitness peak in all environments, explaining its strong overall robustness. Conversely, CEN.PK113-7D and S288C showed no significant changes in robustness but increased fitness by the end of evolution. This outcome could be explained by these populations navigating toward different fitness peaks and achieving higher yet widely distinct fitness in various environments (peaks of different heights), indicative of reduced robustness (light violet planes at different levels in Figure 4.4). The use of fitness landscapes to explain fitness and robustness accounts also for the initially greater robustness of the S288C strain, attributed to its genotypes starting from a valley on the fitness landscape in all environments (same fitness level).

To conclude, fitness-landscape representations can be used to explain evolution trajectories that favor the emergence of robust phenotypes. Yet, they cannot be constructed by relying uniquely on data from **Paper V** and without lineage tracking during evolution.



4

Figure 4.4: Fitness landscapes and robustness. *The top panels report the fitness landscape for Ethanol Red. The x- and y-axes represent the genomic sequence space of all possible genotypes. The z-axis indicates the fitness level for each genotype. The depicted landscape is one of many possible configurations, characterized by varying peaks and valleys. The black dot pinpoints the genotype's location at a selected evolutionary stage. The arrows show the evolutionary trajectory of this genotype within the given environment. The plane serves to illustrate the fitness threshold at which the black dot is positioned. The bottom panels report the fitness landscape for S288C or CEN.PK113-7D, following the same conventions as for Ethanol Red. This landscape visualization aids in understanding the evolutionary path and fitness level of a genotype at a specific point in its evolution.*

Monitoring of evolutionary progress in **Paper V** could have shed light on the genesis of robustness mechanisms. However, this type of analysis was not feasible, as only end-stage evolution samples were phenotyped and sequenced. Revisiting **section 4.1.2**, a subset of the YKO collection was assessed at five different evolutionary time points (114) (**section 2.4:** Dataset 7). Results indicated that evolution in constant environments tended to improve fitness, as reflected by both generation time and yield (Figure 4.5). The only exception was medium containing 50 µM $CdCl_2$, whereby fitness gains were not significant.

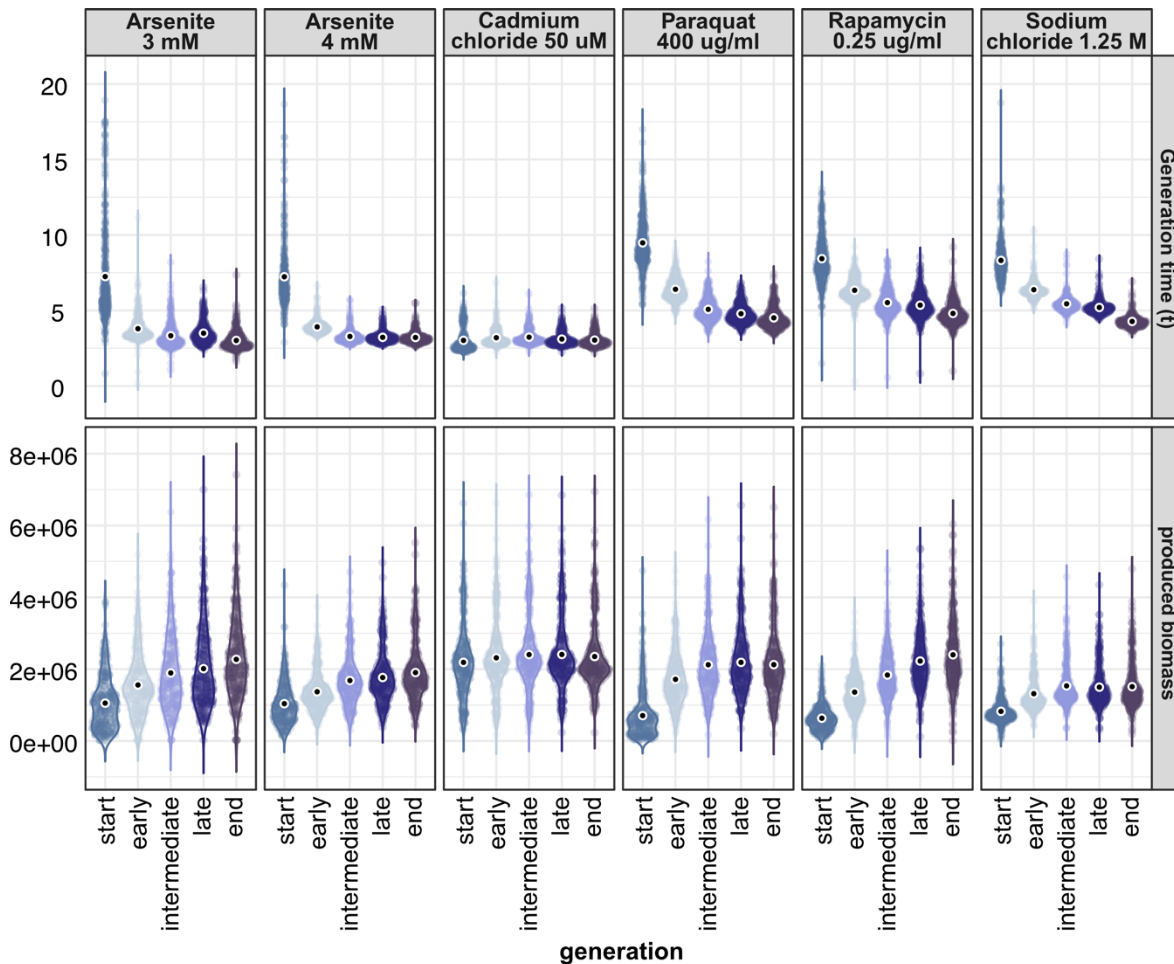


Figure 4.5: Fitness of the YKO collection (Dataset 7) during evolution in constant environments. *Generation time and produced biomass are shown in the two rows (y-axis). Columns denote the media employed in the evolution experiment, while the x-axis reports the time points during evolution (0, 25, 50, 75, and 100 generations). The black point in the middle of the distribution represents the mean fitness across all deletion strains.*

Robustness of the YKO collection (dataset 7) was calculated for each evolution time point, across six environments, and for both generation time and biomass. Robustness increased gradually during evolution, concurrently with fitness improvements (Figure 4.6). Trends in robustness are highly dependent on the evolution set-up and the tested environments. In **Paper V**, Ethanol Red may have experienced a progression similar to the one depicted in Figure 4.6. In contrast, S288C showed no increase in robustness, possibly because the evolutionary environments induced uneven fitness gains (especially under constant conditions).

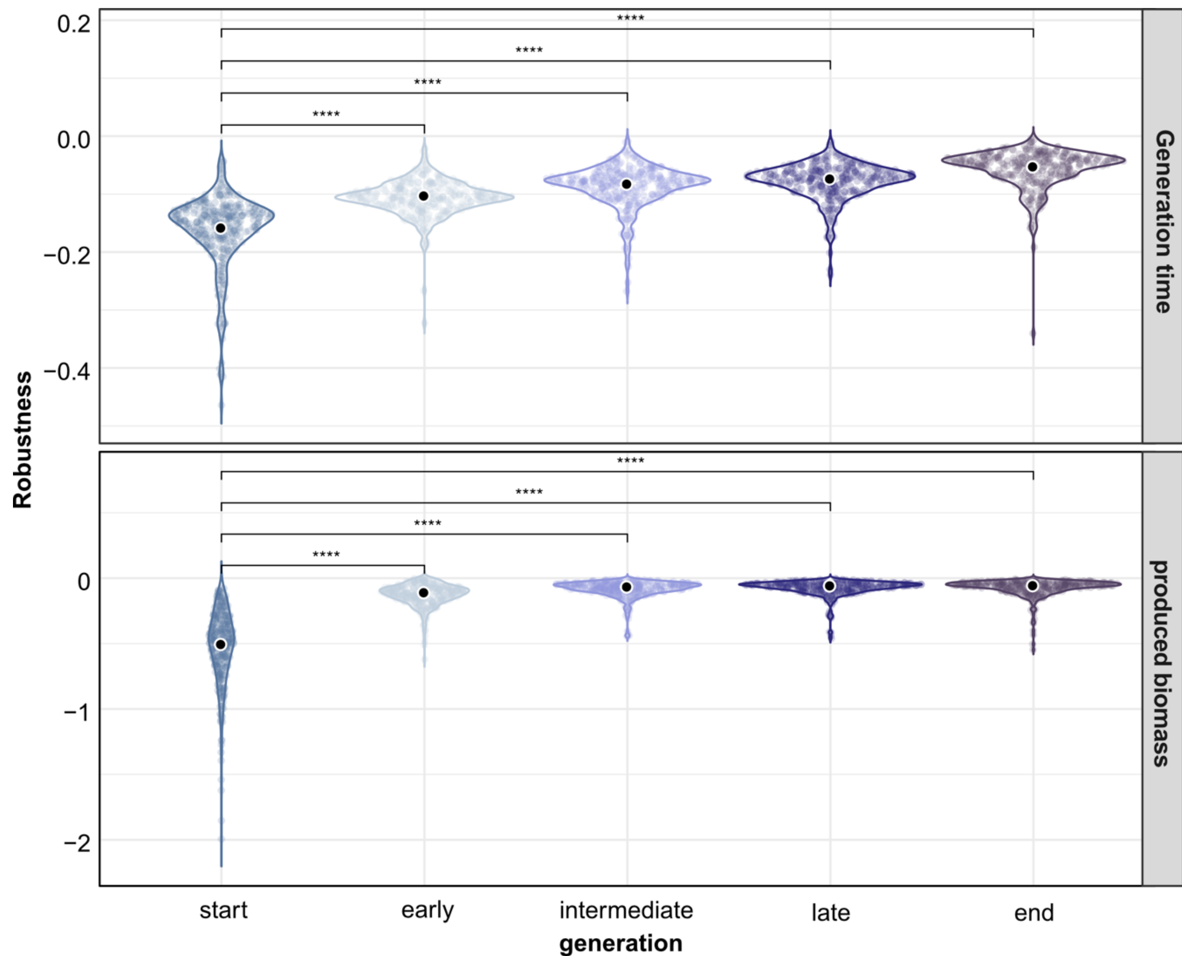


Figure 4.6: Robustness of the YKO collection (Dataset 7) during evolution across six environments**. Robustness for each time point was calculated with Equation 3. Robustness of the generation time and produced biomass is shown in the two rows. Different generations (0, 25, 50, 75, and 100 generations) are reported on the x-axis. The black dots represent the mean of the distributions. Significant differences are indicated by asterisks top of the plots, * (p-value < 0.0001).*

The evolutionary trajectory of each population may vary significantly in fluctuating environments (Figure 4.4). A daily shift in conditions could disrupt progression towards a fitness peak, potentially reverting the population to a lower fitness valley (Figure 4.7). Each environment presents a unique fitness landscape (ENV1 versus ENV2), where a population ascending towards a peak in one condition could find itself descending in another. Adaptations beneficial in ENV1 may lead to a less fit genotype in ENV2, when cells are transferred to a new medium. The sequence of encountered environments shapes the overall fitness of a population (**Paper V**: Figure 3). Therefore, robustness could result from cells being stuck in a valley at the same fitness level due to a changing environment.

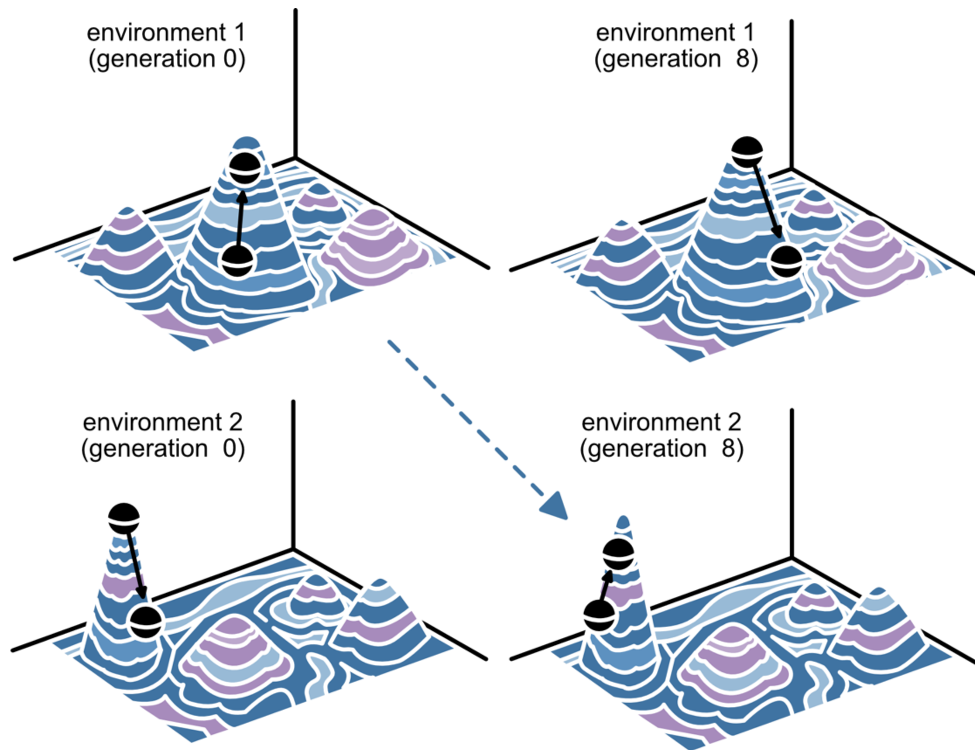


Figure 4.7: Adaptive walks during evolution in fluctuating environments. *All possible genotypes are represented by the x-y plane while fitness is shown on the z-axis. Environments are separated across rows and time during evolution across columns. The dot represents the genotype at a specific time during evolution (generation 0 vs generation 8). The direction of the white arrows denotes the direction the genotype is taking in that specific environment/landscape. The dashed blue arrow shows the transfer of populations to a new environment.*

Building upon the robustness markers discussed in **section 4.1**, the evolutionary experiment and genome sequencing facilitated the analysis of genes that underwent mutations throughout the evolutionary process in both stable and variable environments.

4.2.4 Genomic variants

In **Paper V**, several genes mutated across different populations from diverse evolutionary set-ups and strains were reported (multi-hit genes), suggesting similar selection dynamics in separate populations. Mutations in RHO5 and SNF2 (Table 4.1) had a high impact on CEN.PK113-7D and Ethanol Red evolved in fluctuating environments. Hsp104p, collaborates with Ssa1p and Hsp82p, was also found among the multi-hit genes. Hsp104p has been shown to interact with the yeast prion-like element [PSI+], considered an evolutionary and phenotypic capacitor (193), (194). Prions are nonchromosomal genes composed of altered forms of proteins which can alter other proteins in the same manner (195). Deletion of HSP104 in [PSI+] yeast can irreversibly eliminate [PSI+], thereby disrupting nonsense suppression and potentially altering protein function (194). Here, as

HSP104 mutations were predominantly missense, making the implications for protein function are less clear. Hsp104p interact with yeast prion [PSI+]. The mutated genes from **Paper V** thus relate to the robustness markers discussed in **section 4.1**.

*Table 4.1 Mutated genes identified in the evolution experiment (**Paper V**)*

Gene name	Description	Reference
RHO5	Rho5, a GTPase, is involved in oxidative stress response and apoptosis, with its deletion mutants exhibiting resistance to oxidation	(196)
SNF2	Snf2, a component of the SWI/SNF complex, is essential for chromatin remodeling	(197)
HSP104	Molecular chaperone that clears aggregates after heat shock and propagation of prions	(198)
SSA1	Members of one of the yeast cytosolic hsp70 subfamilies, is involved in translocation of secretory proteins into the lumen of the endoplasmic reticulum	(176)
HSP82	Hsp90 chaperone	(157)

To further validate the implications of genetic mutations on measured fitness and robustness, several methodologies can be used. Genome-wide association studies and quantitative trait loci are the most used approaches (199,200). While these methods were not employed in this **thesis**, an analysis of fitness and robustness among colonies with multiple genetic mutations revealed specific trends (**section 2.4:** Dataset 6). Colonies exhibiting differences in fitness and robustness were picked (Figure 4.8). Those harboring a deletion in PCL6 displayed improved fitness in the presence of HMF; whereas those with mutations in SEC31 and SNF2 were associated with greater fitness under acidic conditions (Figure 4.8.a). Similarly, Ethanol Red colonies carrying mutations in CEP3, IRA2, POL2, SEC31, SNF2, and STB5 exhibited notably higher robustness than the parental strain (Figure 4.8.b). Importantly, these observations do not imply that mutations in these genes are the direct cause of increased fitness and robustness. To establish a causal relationship, genome-wide association studies would be necessary. Additionally, the impact of these mutations could be further verified by introducing them into the ancestral strain and evaluating their effects under the same experimental conditions.

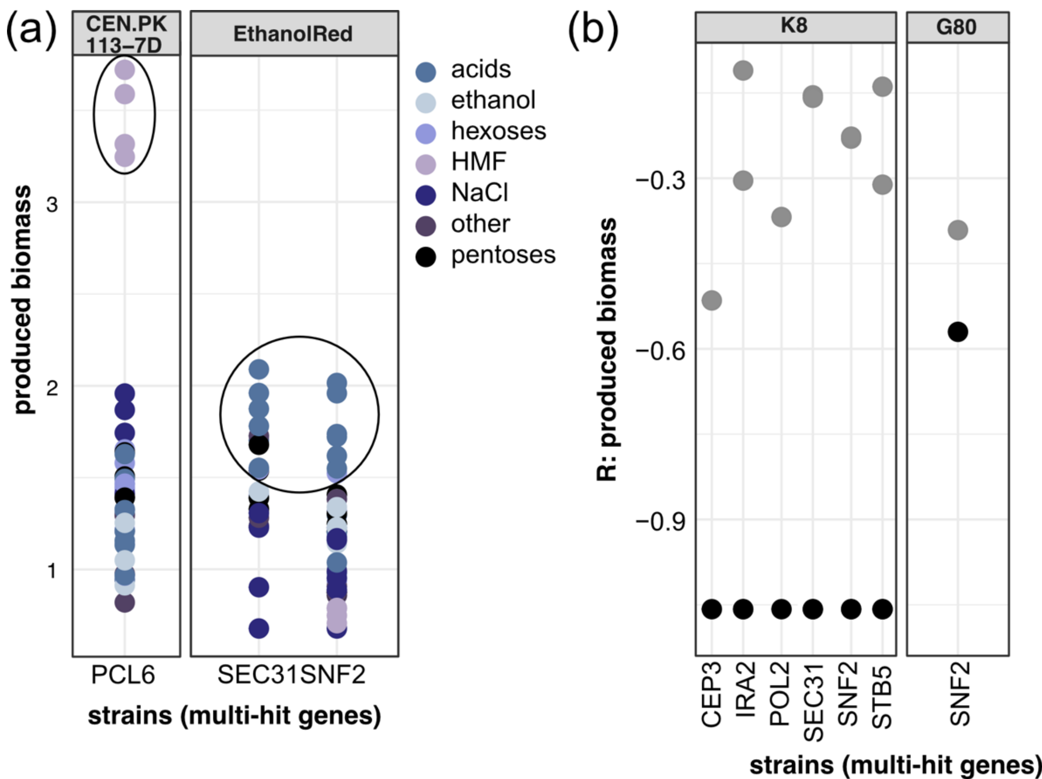


***Figure 4.8: Fitness and robustness of strains carrying mutations in multi-hit genes (Dataset 6)**. a) Three multi-hit genes are shown on the x-axis and the ratio of produced biomass of the evolved samples over that of the parental samples is shown on the y-axis. Strains are reported on top. Colors correspond to different environments. The two circles highlight the highest produced biomass. b) Six genes are listed on the x-axis and robustness of produced biomass is plotted on the y-axis. Black dots correspond to robustness of the parental strain. K8: constant conditions; G80: evolution in fluctuating conditions (80 generations media change).*

4.2.5 Robustness and history-dependent behavior

The concluding part of this **section** on evolutionary dynamics focuses on a non-genetic adaptation method known as acclimatization or short-term adaptation. Populations previously exposed to specific conditions often exhibit reduced lag phase or elevated specific growth rate upon re-exposure to the same condition. This phenomenon is described as history-dependent behavior. Unlike genetic mutations, history-dependent behavior arises from hereditary mechanisms distinct from DNA sequences and is common in yeast adapting to shifts in carbon source (201). Possible explanations for this behavior include epigenetic factors such as DNA methylation, the persistence of proteins and transcription factors that are not broken down following environmental changes, and the cellular metabolic state (202). These mechanisms can explain the observed increase in specific growth rates after transfer to the same medium (**Paper III:** Figure 4).

The results of **Chapter 4** underscore the benefit of investigating robustness mechanisms through evolutionary studies, as well as the use of yeast collections with comprehensive genetic and phenotypic data.

Main Points from Chapter 4
• Hsp90 has been suggested as phenotypic capacitor suppressing phenotypic variation. • Robustness quantification can be applied to yeast deletion collection screens to identify robustness markers. • Sulfur metabolism, heat shock proteins, and ribosomal proteins have emerged as the likeliest effectors of robustness. • Evolution in fluctuating environments enhanced the robustness of the yeast strain Ethanol Red but not the one of laboratory strains CEN.PK113-7D and S288C. The opposite trend was observed for fitness. • Fitness landscapes are useful tools to explain the fitness and robustness trends during evolution. • Mutations in heat shock proteins and in proteins interacting with the yeast prion [PSI+] were hit multiple times during the evolution experiment in both constant and fluctuating conditions.

Chapter 5. Conclusions and Outlook

The work presented in this thesis focuses on the quantification of microbial robustness and exploration of its use in different applications. The first **section** of this chapter outlines how the research objectives presented in **Chapter 1** were achieved. The second **section** presents future perspectives on the presented topics.

5.1 Conclusions

Which quantification theory is better suited to quantify microbial robustness? Which experimental setup can be used such quantification?

To address these questions, different quantification theories were assessed, including Kitano's formula, CV, and the Fano factor, the latter two being indicators of variation. The first two options did not meet some of the criteria required for quantifying robustness according to the definition given in **Review Paper I**. For one, they did not represent measurements of robustness across various phenotypes and scales without depending on subjective choices for controls. In **Paper II**, I showed that the robustness quantification formula based on the normalised Fano factor accurately depicted robustness using experimental data. A normalization factor and a negative sign were added to the formula to facilitate interpretation and to be able to compare the robustness of different phenotypes.

To assess how robust a system's phenotype is when faced with different perturbations, I designed an experimental set-up based on the collection of extensive phenotypic data from numerous strains. This resulted effective when calculating robustness with the formula presented in **Paper II.** The experiment involved simplifying complex perturbation spaces, such as lignocellulose hydrolysates, by applying individual perturbations and then growing strains under these specific conditions.

The combination of experimental design and the robustness quantification method allowed the identification of new phenotypic responses, such as weak acids affecting more significantly ethanol yield than the specific growth rate. This approach also highlighted strains such as Ethanol Red, which possessed robust phenotypes. Additionally, it demonstrated that the expected trade-offs between performance and robustness in yeast did not apply universally across all phenotypes, as happened when evaluating the maximum specific growth rate (**Paper III**).

How can the effect of perturbations be included in a microbial robustness measure?

Perturbations play a crucial role in understanding robustness, and their nature is intricately tied to the nature of robustness. First, in **Papers III and IV,** I demonstrated that robustness was inherently connected to the perturbation space. Various perturbation

spaces, each made up of different individual perturbations, resulted in completely different robustness scores for the same phenotype of a given strain. This insight is particularly critical when evaluating strains for industrial use, where a comprehensive assessment of all potential predictable and stochastic perturbations is essential for a precise measurement of robustness. **Paper III** also indicated that some perturbations did not impact the robustness value in a perturbation space, but analysis of performance in many of them could still be convenient and informative in different processes and other perturbation spaces.

In this **thesis**, I presented three crucial factors to consider when evaluating robustness and indirectly the perturbation space:

I. perturbation probability

II. perturbation intensity

III. interaction effects between perturbations.

These three aspects are not included in the quantification formula so far, but they represent better the industrial environments. Additionally, perturbations are important in the context of evolution as different selective pressures might favour certain populations. It is also worth noting, as demonstrated in **Paper IV**, that some strains exhibit a high level of inherent robustness, which may be largely unaffected by the nature of the perturbation space. However, numerous theoretical works, including those by Kitano, suggest that the concept of a universally robust strain is improbable, particularly due to the existence of trade-offs.

How can adaptive laboratory evolution be used to evolve robustness?

Leveraging the robustness definition from **Review Paper I**, I designed an evolution experiment under fluctuating conditions which, after 300 generations, led to an increased robustness in the industrial strain Ethanol Red (**Paper V**). By evolving strains with different genetic backgrounds, fitness, and ploidy, I found that the inherent characteristics of the strains had a more significant impact on the evolved robustness than the evolutionary regime applied. Specifically, strains with lower initial fitness, such as CEN.PK113-7D and S288C, strongly improved in a fluctuating environment; yet. their robustness levels after evolution remained unchanged after evolution. Fitness landscapes offer a framework to understand how the genetic background influence the evolution of robustness. Still, further experiments are needed to confirm the proposed adaptive walks and especially the role of mutations on both fitness and robustness.

How can genetic and metabolic markers of robustness be identified?

The work presented in this **thesis** revolves primarily around the measurement of phenotypic data within various perturbation spaces for subsequent robustness assessment. However, the quantification formula can be applied to evaluate robustness from a cellular

and metabolic standpoint. In **Papers IV and V,** I applied the robustness quantification formula from **Paper II** to candidate genetic markers or metabolic routes associated with robustness. One approach involved applying the robustness formula to existing datasets that included phenotypic measurements from the yeast deletion collection. Another strategy was to investigate the genetic variants of populations that were associated with increased robustness after evolution in fluctuating environments.

Through this analysis, proteins associated with sulfur metabolism, heat shock proteins, and ribosomal proteins were identified as potential robustness markers, particularly due to their interactions with known phenotypic capacitors such as Hsp90.

In conclusion, the research questions addressed in this **thesis** achieved the goal of establishing a methodology for quantifying microbial robustness, particularly through the use of experimental data, and demonstrated its application across various studies.

5.2 Outlook and Future perspectives

This **thesis** contributes to a greater understanding of microbial robustness. Nevertheless, the present work has also left some questions unanswered, along with new ones that have merged.

5.2.1 Microbial robustness quantification

The robustness quantification formula has been utilized to measure the stability of various phenotypes across several contexts, ranging from specific growth rates to product yields. The same equation was applied also in a different study to evaluate the robustness of intracellular parameters monitored through biosensors, such as intracellular ATP levels. This involved monitoring the stability of phenotypes over time and measuring population heterogeneity using single-cell analyses in microfluidic devices (126,127). Hence, the formula has proven to be widely applicable in diverse contexts.

While the robustness quantification formula introduced in **Paper II** has proven helpful, it requires further refinement and development. Currently, assessment of the perturbation space relies on review of available literature or consultation with field experts which are heavily influenced by individual experience, highlighting the need for a more systematic method that selects perturbations based on empirical data. Additionally, the three perturbation aspects—combined effects, intensity, and probability—discussed in **Chapter 2,** are not effectively incorporated into the quantification formula. Gaining access to actual measurements from large-scale fermenters, especially from a real industrial site, would improve our understanding of the fluctuations that occur in such settings. This could be achieved by analyzing data from available measurements such as temperature or pH

sensors, both of which have been mentioned in the survey as significant contributors to variability. Alternatively, one could implement flow-following sensor devices which describe gradients in the reactors (203), to provide better guidance on the types of fluctuations, their probability, and their intensity in industrial fermenters. Computational fluid dynamic models have also great potential in predicting gradients and perturbations in bioreactors (204). Integration of real perturbation data on the robustness quantification would allow better reproducibility of the performance metric measured at lab-scale in large-scale processes.

Another limitation hindering the routine implementation of robustness measurements is the lack of high-throughput equipment for phenotypic measurements. The experimental techniques employed in this **thesis**, such as the growth profiler, scan-o-matic, and Adaptive Laboratory Evolution (ALE) in 96-well plates outlined in **Chapter 2**, enabled the rapid collection of extensive phenotypic data, which would not have been feasible with basic cultivation methods, such as plate readers or shake flasks. However, high-throughput screenings present their own set of challenges. They do not accommodate certain perturbations, such as those involving high pressure, strictly anaerobic conditions or temperatures other than 30°C. Additionally, high-throughput methods can make it difficult to measure certain phenotypes such as cell dry weight, and collecting samples during the growth process is not feasible, complicating the measurement of ethanol productivity.

Liquid handlers simplify the process by enabling the rapid dispensing of substantial volumes of media and facilitating culture transfers, thus minimizing experimental timelines. Liquid chromatography equipped with 96-well plate configurations can be used to determine the concentrations of sugars and inhibitors on a large scale. For a more accurate simulation of anaerobic conditions or varied bioreactor setups, the BioLector (Beckman) offers a preferable alternative to the growth profiler. Nonetheless, the ideal system that fulfils all these requirements, particularly one that is fully automated, has yet to be realized. Compromises among accuracy, sensitivity, speed, and high throughput remain inevitable.

Furthermore, the robustness quantification method described in **Paper II** does not distinguish for high or low performing strains. Consequently, strains with low fitness may still exhibit high robustness. While this can be valuable for physiological studies, it may not be desirable when developing industrial strains. For future robustness assessments in bioprocesses, it would be beneficial to weigh performance and fitness equally. Incorporating Pareto efficiency into the evaluation of robustness and fitness, as suggested in **Chapter 3**, could aid in identifying strains that optimally balance high fitness and robustness.

An additional point to consider for future research is that when different perturbations are tested, the actual cultivation might evaluate different subpopulations, even if cultures originate from the same initial population. Past research has shown that these subpopulations, and even individual cells, can exhibit very different behavior within the same overall population (205). To address population heterogeneity concerns, robustness

should be evaluated by exposing a population to a series of perturbations in bioreactors. This approach may face challenges, such as the appearance of mutations in the population. One intriguing option would be to investigate the robustness of the transcriptome during a chemostat cultivation, when the culture is subjected to various perturbations. With technological progress in reactor design, I anticipate robustness quantification to be feasible on a scale larger than the current 250-µl cultures in 96-well plates, and rely instead on Pioreactors (pioreactor ©, Canada) with 20-mL volumes.

Addressing the points described above regarding perturbations, performance, and heterogeneity could significantly enhance the integration of robustness quantification in the design of microbial strains, making it an essential step prior to scale-up. Companies such as Genomatica are already implementing a similar approach using small-scale reactors during the pre-scale-up stage. As emerged from the survey conducted in this thesis, business developers, R&D directors, and process scientists favor the development of robust cell factories more than process control. This preference stems from the fact that the latter is expensive and require continuous maintenance.

5.2.2 Metabolic and genetic markers of robustness

Based on the findings from **Paper IV and V**, as well as the cumulative insights from this **thesis**, my opinion lies more with pursuing robust design strategies over searching for specific robustness markers. While there are promising associations between certain cellular processes and robustness, it is clear that this area of research is still in its infancy. Additionally, it remains unclear whether these markers are unique to certain perturbation spaces or if they have a universal value. Insights from **Paper IV** suggest a dual nature. Deletion of MET28 appears to confer general robustness to the specific growth rate of the CEN.PK113-7D strain; whereas other markers such as SMA2, seem to be perturbation space-specific. Reflecting on **Chapter 3** and the discussions on the sub-features of robustness, it is evident that a phenotype's external robustness does not necessarily imply stability of its internal cellular features. In fact, it may be quite the contrary; to preserve phenotypic stability, internal cellular features might have to adapt or change.

The theories surrounding phenotypic and evolutionary capacitors that buffer for phenotypic variation are certainly valuable, but more should be done to determine the context in which these capacitors are functional compared to other genes. The question remains as to why certain strains exhibit greater robustness in relation to specific phenotypes. Based on current data, it would be interesting to further investigate areas, such as protein degradation and regulation, phenotypic capacitors, and ploidy levels.

In the absence of a clearly defined methodology, designing a robust strain remains a complex task. Adaptive evolution in fluctuating environments appears to be a promising

approach, but as shown in **Paper V**, its success is strain-specific, and the link to ploidy and fitness requires more validation. Deletion of MET28 is also of interest, though it significantly reduced strain fitness, indicating a need for further tests across different strains.

Future work should include more comprehensive analysis of yeast deletion collections to evaluate robustness across various conditions and phenotypes. It is also important to assess the robustness of the transcriptome under different environmental challenges. Additionally, leveraging genome-scale metabolic models to target phenotypic robustness while controlling gene deletions or perturbing the metabolic networks could provide insights into designing robust strains.

Over time, research efforts such as those presented in this **thesis** and the methodologies proposed in this discussion are expected to offer strategies for the engineering of strains with inherently robust traits.

Acknowledgements

This acknowledgment page represents the end of this PhD journey, that you can now see just in form of scientific work and graphs and data, but it has been so much more. The PhD was not only intellectual and academic growth, but it really was, simply, personal growth. I really think that the major contributor of this growth is all of you. My grandmother used to say that if you have your hand closed nothing can go out, but nothing can get in either. And I personally think this well describe my journey. One can be so stuck in its own beliefs and personal comfort when it comes to knowledge. Studying something completely new is very scary! But being able to overcome the fear and opening that hand, brings satisfaction and pride.

First, I want to thank my supervisor, Lisbeth. When I started working in IndBio, I was a biotech baby, I barely knew how to grow yeast. You have really supported me not only scientifically, but also understanding my space, my needs and took in all the crazy ideas and opportunities I proposed you with enthusiasm. You have encouraged me through my insecurities, and you have really built an environment where people and I, personally, could thrive and develop. Thank you for your great help and support, especially when writing my thesis, and most of all thank you for setting a great example for women in science.

I want to thank my examiner, Calle. You were the first person that sparked my interest in industrial biotechnology. Talking with you about science, my research and simply life, has contributed significantly to my development and made me feel like my ideas had value in this scary research world. Thank you for taking the time in ensuring everybody is heard.

This work would not have been possible without the significant contributions of the Robustness Team (Lisbeth, Luca, Peter and Nathalia). This topic was so vast to navigate but thank to you, my interest grew significantly, and I have learnt how to approach science with a more critical eye. A special thanks to Nathalia, that in this last year has been supporting me in so many ways. Taking the lead of not only research projects but also contributing to creating a special place in our office and supporting me in general. I want to thank ALTAR (now Ginkgo Bioworks) as well, for the project on evolution, it has been great to collaborate with you guys, even if the project was not included in the thesis it is now in very good hands (wink wink Nathalia).

I want to give a special thank to Michael and his group. You welcomed me in your group and, from day one, I felt like I have always been part of it. I have learnt so much from you and everyone else in the group. Your approach to science is really different from what I was used to. Being an engineer, I have always limited my research to its application, but you have showed me how important it is to dive into the theoretical aspects of biology, evolution and science in general. I'm very happy for how the project has developed, and I

deeply appreciate your solid support and guidance! I here want to mention Artur, Misha, and Thomas without you, I'd still be figuring out genomic DNA pools and programming of the Biomek. Your guidance has been indispensable! In this section I also want to mention Eliot, Josh, Stephanie, Adriana, Ben and Clare. You guys adopted me, and I will never forget the fun times we had together. You are great!

Paper V would not have been finished without the great help of Karl, Diana and Estelle. Thank to Diana and Estelle to deal with my complicated data and supporting me with the WGS analysis. I have learnt a lot from you! Karl, you are an amazing collaborator. Not only you are very knowledgeable, but you also support your students in a very humble way. Thank you for all the help with scan-o-matic and corrections in general. It was so fun to run experiments with you! I just wish I could do my PhD on the honey yeast with you now!

I want to thank Vijay. I would not have done my project in US if it wasn't for you. You always find the time to help everyone. Your suggestions and the interesting papers you sent me contributed significantly to my discussions and my thesis. You set a good example for all of us, so thank you for setting up the best research environment!

I want to thank Sofia, my master student, for the work on the combined effects of perturbations. Your enthusiasm and ability to take in information in such short time was impressive! I am sure you will achieve great things and it was a pleasure to collaborate with you.

A BIG THANK to IndBio and all the people passing by. All of you make this place special and let every single person blossom in its own way. Thank you for respecting diversity in all its forms. Thank to Tove and Amanda, FIKA time and our shared office made my work environment really sweet! Long live Giovanna!

I am afraid this is only part of my thank-yous. I need to make room to give a big thanks to everyone who has been in my life, whether it's been for a while or just recently, for all their solid support.

First of all, Elena. Since day 1 you have always been there. It didn't take so much time for me to become the COZZA in your life. There is not a person that has supported me more than you in these years. You were close to me in many up and downs. Elena, you have always been an example for me of a person that pursue her goals with decision and determination. I hope I have 1/10 of your strength in life to achieve what I dream but I am sure that in the meantime you will be my 9/10. Thank you for being such an amazing and attentive friend.

Luca. Now, after your defense, it seems I have already told you 700 times how important you have been in these years. "Ce la facciamo, penso" and at the end I think we really have done it. You have pushed me to achieve things in the best possible way, and you have taught me so much! Thank you for all the help throughout the years and to be

supporter number one of my quantification formula. We have supported each other continuously and for this I think our two PhD really complemented each other. We can't predict the future, but I bet if I stumble upon a well-organized lab, you've probably left your mark there. A fine anno, alla fine, ci siamo arrivati. Bird noises for you.

Ibai. You know that feeling when you picture a peaceful place to relax? Well, for me, sometime, it's sitting under a tree in a sunny day. That's exactly how you've made me feel all these years - thanks for bringing warmth and happiness into my days.

Mauro. Our 89% of shared personality made it easy for me to naturally orbit around you. In this last period, our friendship has developed into something deep, and I couldn't ask for a better companion in these last months. You brought back my artistic side, my philosophical thinking and I am sure that side of me that you contributed to, helped me in finishing this thesis and being proud of my work. Thank you for snapping me out of things *WRONG* and for simply being there.

Simone. Vecio, grazie per condividere con me tutti questi momenti di sclero. Il codice, la Svezia, l'Italia, sentirsi senza casa. La vårdcentral, US, la carriera, Massimo Pericolo e la bassa bresciana. Il male all'anca. Sei stato un punto fisso in questi anni e grazie per darmi sempre un punto di vista concreto e ben inquadrato. Ti s(t)imo molto e tante delle mie opinioni sul mondo, la società e la politica sono state influenzate da te.

Qui devo menzionare anche quella friulana di Veronica. Sei pazza, e la tua pazzia mi ha riempita di gioia tantissime volte. Grazie per avermi contagiato e reso partecipe della tua leggerezza.

Maurizio Maria Mormino. MMM. Maria. what can I say here?! Mauri, non hai idea di quanto tu abbia influito su di me in questi anni. Non so bene da quando, siccome ti metto e tolgo dai posti così, randomly. Ogni singola cosa che mi é capitata ha ricevuto i tuoi consigli, il tuo supporto, le tue risate, il tuo conforto in generale. *Ehm Cecia* sei stato un compagno d'ufficio insostituibile. COLLEGA. Posso solo sperare di perdere a pingpong e a ogni altro gioco possibile (tranne Mario Kart) ancora per tanti anni. Grazie per avermi accolto ed essere rimasto nella mia vita. Grazie per avermi mostrato le cose importanti in questi anni. Grazie Mauri.

Un enorme grazie alla mia famiglia. Mia mamma Maria. Grazie per avermi insegnato quanto è importante studiare, avere sempre sostenuto il pensiero indipendente e la ricerca di cose belle e per avermi dato la possibilità di cogliere le opportunità che mi sono state date. Grazie per avermi insegnato il valore della bontà e del buon cuore. Mia sorella Eloisa, che è sempre stata al mio fianco, la mia forza e il mio sostegno e la mia cheerleader numero 1. Senza di te, con la tua innata capacità di capire senza aver bisogno di dire niente non sarei arrivata fino a qui. Grazie a mio padre Claudio che ha sempre stimolato in me la voglia di conoscenza. La tua voglia e capacità di raccontare, mi ha insegnato il valore della

comunicazione e dell’ascolto. Grazie a mio zio Massimo. Con la tua immensa umiltà mi sei sempre stato vicino, affiancandomi con ciò di cui avevo bisogno, anche senza che io lo sapessi. Grazie a Daniela, che da tanti anni è parte della mia famiglia. La tua premura, attenzione ai bisogni, la tua forza e ricerca di felicità mi hanno sempre spronato a raggiungere i miei traguardi e allo stesso tempo a non lasciare nessuno indietro.

Grazie alle Buddine. Paola, ma che ti devo dire che già non sai, che abbiamo pettinato le Barbie insieme da quando avevamo 3 anni? Sei il mio porto sicuro. Forse quel poco di casa che ancora sento. Grazie per aver sostenuto questo mio viaggio in terre lontane, che dovevo andare all’estero a divenire una scienziata. Irene, inutile dire che in questi ultimi anni la tua presenza é stata fondamentale. Sei una donna incredibile e la tua forza e voglia di raggiungere sempre il meglio mi trasmette energia anche da distante. Chiara, tu sei esattamente come il Moncenisio, calmo e indescrivibile. Ogni volta che mi son sentita di mollare tutto, immaginare di venire da te, li su in montagna, mi ha sempre rimesso in sesto. La pace che mi fai sentire quando siamo insieme é impagabile. Grazie mie rocce.

L’ultimo, importante, ringraziamento lo devo a TE. **Andrea**. Nel libro Kern’s Process Heat Transfer ero stata profondamente colpita dalla dedica di Donald Kern che diceva: “ To my wife, Natalie, for her real help”. Perché questo é stato. Real Help. Da sempre mi hai spronato a raggiungere i miei obiettivi e a pormene di nuovi. Ma soprattutto mi hai indirizzato a rincorrere ciò a cui do realmente valore. Mi hai seguito nei miei spostamenti e sostenuto nelle mie scelte. Sei stato il più grande contribuente al mio lato critico e pragmatico e non mi stancherò mai dei nostri lunghi discorsi sulla scienza e su tutto ciò che riguarda il mondo. Grazie per il tuo affetto quotidiano e per essere nella mia vita. Sei da sempre la mia calma e il mio lato lucido. Questa tesi, questo lavoro, questi anni, li dedico a te.

“C'è un tempo d'aspetto come dicevo

Qualcosa di buono che verrà”

E infine grazie a quella piccola Cecilia che aveva ritagliato l’articolo di Le Scienze sulle le lumache smeraldine.

References

1. Kitano H. Biological robustness. Nature Reviews Genetics. 2004. p. 826–37.
2. Olsson L, Rugbjerg P, Torello Pianale L, Trivellin C. Robustness: linking strain design to viable bioprocesses. Trends Biotechnol [Internet]. 2022 Feb 1 [cited 2022 Feb 4]; Available from: https://linkinghub.elsevier.com/retrieve/pii/S016777992200004X
3. Wehrs M, Tanjore D, Eng T, Lievense J, Pray TR, Mukhopadhyay A. Engineering Robust Production Microbes for Large-Scale Cultivation. Trends Microbiol [Internet]. 2019 Jun;27(6):524–37. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0966842X19300198
4. Gong Z, Nielsen J, Zhou YJ. Engineering Robustness of Microbial Cell Factories. Biotechnol J [Internet]. 2017 Oct 18 [cited 2021 Aug 23];12(10):1700014. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/biot.201700014
5. Whitacre JM. Biological robustness: Paradigms, mechanisms, systems principles. Front Genet. 2012;3(MAY):67.
6. Stelling J, Sauer U, Szallasi Z, Doyle FJ, Doyle J. Robustness of Cellular Functions. Cell. 2004 Sep 17;118(6):675–85.
7. Félix MAA, Barkoulas M. Pervasive robustness in biological systems. Nature Reviews Genetics 2015 16:8 [Internet]. 2015 Jul 17 [cited 2021 Oct 13];16(8):483–96. Available from: https://www.nature.com/articles/nrg3949
8. Masel J, Siegal ML. Robustness: mechanisms and consequences. 2009 Sep 1 [cited 2022 Mar 4];25(9). Available from: /pmc/articles/PMC2770586/
9. Kitano H. Violations of robustness trade-offs [Internet]. Vol. 6, Molecular Systems Biology. European Molecular Biology Organization; 2010 [cited 2021 May 26]. p. 384. Available from: /pmc/articles/PMC2913402/
10. Mohedano MT, Konzock O, Chen Y. Strategies to increase tolerance and robustness of industrial microorganisms. Synth Syst Biotechnol. 2022 Mar 1;7(1):533–40.
11. Neubauer P, Junne S. Scale-down simulators for metabolic analysis of large-scale bioprocesses. Curr Opin Biotechnol [Internet]. 2010 Feb [cited 2022 Mar 22];21(1):114–21. Available from: https://pubmed.ncbi.nlm.nih.gov/20185293/
12. Torello Pianale L, Rugbjerg P, Olsson L. Real-Time Monitoring of the Yeast Intracellular State During Bioprocesses With a Toolbox of Biosensors. Front Microbiol. 2022 Jan 7;12:802169.
13. Rego-Costa A, Huang IT, Desai MM, Gombert AK. Yeast population dynamics in Brazilian bioethanol production. G3 Genes|Genomes|Genetics [Internet]. 2023 Jul 5 [cited 2024 Jan 20];13(7):104. Available from: https://dx.doi.org/10.1093/g3journal/jkad104
14. Abreu CI, Friedman J, Andersen Woltz VL, Gore J. Mortality causes universal changes in microbial community composition. Nature Communications 2019 10:1 [Internet]. 2019 May 9 [cited 2024 Jan 20];10(1):1–9. Available from: https://www.nature.com/articles/s41467-019-09925-0
15. Kitano H. Towards a theory of biological robustness. Mol Syst Biol [Internet]. 2007 [cited 2022 Mar 4];3:137. Available from: /pmc/articles/PMC2013924/
16. Zhu L, Zhu Y, Zhang Y, Li Y. Engineering the robustness of industrial microbes through synthetic biology. Trends Microbiol [Internet]. 2012 Feb 1 [cited 2021 Feb 25];20(2):94–101. Available from: https://www.sciencedirect.com/science/article/pii/S0966842X11002149
17. Gonzalez R, Morales P, Tronchoni J, Cordero-Bueso G, Vaudano E, Quirós M, et al. New genes involved in osmotic stress tolerance in saccharomyces cerevisiae. Front Microbiol. 2016 Sep 28;7(SEP):218934.
18. van Dijk M, Trollmann I, Saraiva MAF, Brandão RL, Olsson L, Nygård Y. Small scale screening of yeast strains enables high-throughput evaluation of performance in lignocellulose hydrolysates. Bioresour Technol Rep. 2020 Sep 1;11:100532.
19. Thomsen EA, Mikkelsen JG. CRISPR-Based Lentiviral Knockout Libraries for Functional Genomic Screening and Identification of Phenotype-Related Genes. Methods Mol Biol [Internet]. 2019 [cited 2023 Aug 21];1961:343–57. Available from: https://pubmed.ncbi.nlm.nih.gov/30912056/
20. Tesic M, Marsh JA, Cullinan SB, Gaber RF. Functional Interactions between Hsp90 and the Co-chaperones Cns1 and Cpr7 in Saccharomyces cerevisiae. Journal of Biological Chemistry. 2003 Aug 29;278(35):32692–701.
21. Abas N, Kalair A, Khan N. Review of fossil fuels and future energy technologies. Futures. 2015 May 1;69:31–49.

22. Paraschiv S, Paraschiv LS. Trends of carbon dioxide (CO2) emissions from fossil fuels combustion (coal, gas and oil) in the EU member states from 1960 to 2018. Energy Reports. 2020 Dec 1;6:237–42.

23. Ali KA, Ahmad MI, Yusup Y. Issues, Impacts, and Mitigations of Carbon Dioxide Emissions in the Building Sector. Sustainability 2020, Vol 12, Page 7427 [Internet]. 2020 Sep 10 [cited 2024 Jan 28];12(18):7427. Available from: https://www.mdpi.com/2071-1050/12/18/7427/htm

24. Ritchie H, Rosado P, Roser M. Emissions by sector: where do greenhouse gases come from? Our World in Data [Internet]. 2024 Jan 5 [cited 2024 Jan 28]; Available from: https://ourworldindata.org/emissions-by-sector

25. Nationally determined contributions to climate change, World, 2023 [Internet]. [cited 2024 Jan 28]. Available from: https://ourworldindata.org/grapher/nationally-determined-contributions

26. Frey GW, Linke DM. Hydropower as a renewable and sustainable energy resource meeting global energy challenges in a reasonable way. Energy Policy. 2002 Nov 1;30(14):1261–5.

27. Vorobyev R, Metzger B. Hydropower maintaining its supremacy in 2014. Renewable Energy Focus. 2015 Dec 1;16(5–6):135–7.

28. Renewable electricity generation, World [Internet]. [cited 2024 Jan 28]. Available from: https://ourworldindata.org/grapher/modern-renewable-energy-consumption

29. Modern renewable energy generation by source, World [Internet]. [cited 2024 Mar 6]. Available from: https://ourworldindata.org/grapher/modern-renewable-prod

30. Hill J, Nelson E, Tilman D, Polasky S, Tiffany D. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. Proc Natl Acad Sci U S A [Internet]. 2006 Jul 25 [cited 2024 Feb 25];103(30):11206–10. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.0604600103

31. Nass LL, Pereira PAA, Ellis D. Biofuels in Brazil: An Overview. Crop Sci [Internet]. 2007 Nov 1 [cited 2024 Jan 28];47(6):2228–37. Available from: https://onlinelibrary.wiley.com/doi/full/10.2135/cropsci2007.03.0166

32. Trumbo JL, Tonn BE. Biofuels: A sustainable choice for the United States' energy future? Technol Forecast Soc Change. 2016 Mar 1;104:147–61.

33. Clarke KG. Microbiology. Bioprocess Engineering [Internet]. 2013 Jan 1 [cited 2024 Feb 26];7–24. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9781782421672500028

34. Moreno AD, Susmozas A, Oliva JM, Negro MJ. Overview of bio-based industries. Biobased Products and Industries. 2020 Jan 1;1–40.

35. From the sugar platform to biofuels and biochemicals - European Commission [Internet]. [cited 2024 Feb 26]. Available from: https://energy.ec.europa.eu/publications/sugar-platform-biofuels-and-biochemicals_en

36. Siracusa V, Blanco I. Bio-Polyethylene (Bio-PE), Bio-Polypropylene (Bio-PP) and Bio-Poly(ethylene terephthalate) (Bio-PET): Recent Developments in Bio-Based Polymers Analogous to Petroleum-Derived Ones for Packaging and Engineering Applications. Polymers (Basel) [Internet]. 2020 Jul 23 [cited 2024 Feb 26];12(8):1641. Available from: /pmc/articles/PMC7465145/

37. Enriquez E, Mohanty AK, Misra M. Biobased polymer blends of poly(trimethylene terephthalate) and high density polyethylene. Mater Des. 2016 Jan 15;90:984–90.

38. Cho JS, Kim GB, Eun H, Moon CW, Lee SY. Designing Microbial Cell Factories for the Production of Chemicals. JACS Au [Internet]. 2022 Aug 22 [cited 2024 Jan 26];2(8):1781–99. Available from: https://pubs.acs.org/doi/full/10.1021/jacsau.2c00344

39. Calero P, Nikel PI. Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. Microb Biotechnol [Internet]. 2019 Jan 1 [cited 2024 Jan 25];12(1):98. Available from: /pmc/articles/PMC6302729/

40. Toivari MH, Salusjärvi L, Ruohonen L, Penttilä M. Endogenous Xylose Pathway in Saccharomyces cerevisiae. Appl Environ Microbiol [Internet]. 2004 Jun [cited 2024 Jan 28];70(6):3681. Available from: /pmc/articles/PMC427740/

41. Markina NM, Kotlobay AA, Tsarkova AS. Heterologous Metabolic Pathways: Strategies for Optimal Expression in Eukaryotic Hosts. Acta Naturae [Internet]. 2020 Aug 7 [cited 2024 Jan 28];12(2):28. Available from: /pmc/articles/PMC7385092/

42. Preiss L, Hicks DB, Suzuki S, Meier T, Krulwich TA. Alkaliphilic Bacteria with Impact on Industrial Applications, Concepts of Early Life Forms, and Bioenergetics of ATP Synthesis. Front Bioeng Biotechnol [Internet]. 2015 [cited 2024 Jan 25];3(JUN). Available from: /pmc/articles/PMC4453477/

43. Thorwall S, Schwartz C, Chartron JW, Wheeldon I. Stress-tolerant non-conventional microbes enable next-generation chemical biosynthesis. Nat Chem Biol [Internet]. 2020 Feb 1 [cited 2024 Jan 25];16(2):113–21. Available from: https://pubmed.ncbi.nlm.nih.gov/31974527/

44. Crosby JR, Laemthong T, Lewis AM, Straub CT, Adams MW, Kelly RM. Extreme thermophiles as emerging metabolic engineering platforms. Curr Opin Biotechnol [Internet]. 2019 Oct 1 [cited 2024 Jan 25];59:55–64. Available from: https://pubmed.ncbi.nlm.nih.gov/30875665/

45. Yang Z, Leero DD, Yin C, Yang L, Zhu L, Zhu Z, et al. Clostridium as microbial cell factory to enable the sustainable utilization of three generations of feedstocks. Bioresour Technol. 2022 Oct 1;361:127656.

46. Hwang S, Joung C, Kim W, Palsson B, Cho BK. Recent advances in non-model bacterial chassis construction. Curr Opin Syst Biol. 2023 Dec 1;36:100471.

47. Jiang W, Gao H, Sun J, Yang X, Jiang Y, Zhang W, et al. Current status, challenges and prospects for lignin valorization by using Rhodococcus sp. Biotechnol Adv. 2022 Nov 1;60:108004.

48. Sun L, Alper HS. Non-conventional hosts for the production of fuels and chemicals. Curr Opin Chem Biol. 2020 Dec 1;59:15–22.

49. Rebello S, Abraham A, Madhavan A, Sindhu R, Binod P, Karthika Bahuleyan A, et al. Non-conventional yeast cell factories for sustainable bioprocesses. FEMS Microbiol Lett [Internet]. 2018 Nov 1 [cited 2024 Jan 28];365(21):222. Available from: https://dx.doi.org/10.1093/femsle/fny222

50. Parapouli M, Vasileiadis A, Afendra AS, Hatziloukas E. Saccharomyces cerevisiae and its industrial applications. AIMS Microbiol [Internet]. 2020 [cited 2024 Jan 29];6(1):1. Available from: /pmc/articles/PMC7099199/

51. Sharma J, Kumar V, Prasad R, Gaur NA. Engineering of Saccharomyces cerevisiae as a consolidated bioprocessing host to produce cellulosic ethanol: Recent advancements and current challenges. Biotechnol Adv. 2022 May 1;56:107925.

52. Beran D, Laing RO, Kaplan W, Knox R, Sharma A, Wirtz VJ, et al. A perspective on global access to insulin: a descriptive study of the market, trade flows and prices. Diabet Med [Internet]. 2019 Jun 1 [cited 2024 Jan 25];36(6):726–33. Available from: https://pubmed.ncbi.nlm.nih.gov/30888075/

53. Vanderwaeren L, Dok R, Voordeckers K, Nuyts S, Verstrepen KJ. Saccharomyces cerevisiae as a Model System for Eukaryotic Cell Biology, from Cell Cycle Control to DNA Damage Response. International Journal of Molecular Sciences 2022, Vol 23, Page 11665 [Internet]. 2022 Oct 1 [cited 2024 Jan 15];23(19):11665. Available from: https://www.mdpi.com/1422-0067/23/19/11665/htm

54. Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, et al. Population genomics of domestic and wild yeasts. Nature 2009 458:7236 [Internet]. 2009 Feb 11 [cited 2024 Jan 15];458(7236):337–41. Available from: https://www.nature.com/articles/nature07743

55. Strope PK, Skelly DA, Kozmin SG, Mahadevan G, Stone EA, Magwene PM, et al. The 100-genomes strains, an S. cerevisiae resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. Genome Res [Internet]. 2015 May 1 [cited 2024 Jan 25];25(5):762. Available from: /pmc/articles/PMC4417123/

56. Peter J, De Chiara M, Friedrich A, Yue JX, Pflieger D, Bergström A, et al. Genome evolution across 1,011 Saccharomyces cerevisiae isolates. Nature 2018 556:7701 [Internet]. 2018 Apr 11 [cited 2024 Jan 29];556(7701):339–44. Available from: https://www.nature.com/articles/s41586-018-0030-5

57. Kong II, Turner TL, Kim H, Kim SR, Jin YS. Phenotypic evaluation and characterization of 21 industrial Saccharomyces cerevisiae yeast strains. FEMS Yeast Res [Internet]. 2018 Feb 1 [cited 2024 Jan 25];18(1):1. Available from: https://dx.doi.org/10.1093/femsyr/foy001

58. da Conceição LEFR, Saraiva MAF, Diniz RHS, Oliveira J, Barbosa GD, Alvarez F, et al. Biotechnological potential of yeast isolates from cachaça: the Brazilian spirit. J Ind Microbiol Biotechnol [Internet]. 2015 Feb 1 [cited 2021 Nov 4];42(2):237–46. Available from: https://academic.oup.com/jimb/article/42/2/237/5995451

59. Rogers DF, Shafer SM. Measuring cellular manufacturing performance. Manufacturing Research and Technology. 1995 Jan 1;24(C):147–65.

60. Meissner MP, Woodley JM. Mass-based biocatalyst metrics to guide protein engineering and bioprocess development. Nature Catalysis 2022 5:1 [Internet]. 2022 Jan 28 [cited 2024 Jan 26];5(1):2–4. Available from: https://www.nature.com/articles/s41929-021-00728-5

61. Van Dien S. From the first drop to the first truckload: commercialization of microbial processes for renewable chemicals. Curr Opin Biotechnol. 2013 Dec 1;24(6):1061–8.

62. Han Y, Zhang F. Control strategies to manage trade-offs during microbial production. Curr Opin Biotechnol. 2020 Dec 1;66:158–64.

63. Rodríguez TH, Morerod S, Pörtner R, Wurm FM, Frahm B. Considerations of the Impacts of Cell-Specific Growth and Production Rate on Clone Selection—A Simulation Study. Processes 2021, Vol 9, Page 964 [Internet]. 2021 May 28 [cited 2024 Jan 26];9(6):964. Available from: https://www.mdpi.com/2227-9717/9/6/964/htm

64. Liu Y, el Masoudi A, Pronk JT, van Gulik WM. Quantitative Physiology of Non-Energy-Limited Retentostat Cultures of Saccharomyces cerevisiae at Near-Zero Specific Growth Rates. Appl Environ Microbiol [Internet]. 2019 [cited 2024 Feb 26];85(20). Available from: https://pubmed.ncbi.nlm.nih.gov/31375494/

65. Cabrera E, Welch LC, Robinson MR, Sturgeon CM, Crow MM, Segarra VA. Cryopreservation and the Freeze–Thaw Stress Response in Yeast. Genes (Basel) [Internet]. 2020 Aug 1 [cited 2024 Jan 26];11(8):1–11. Available from: /pmc/articles/PMC7463829/

66. Della-Bianca BE, Basso TO, Stambuk BU, Basso LC, Gombert AK. What do we know about the yeast strains from the Brazilian fuel ethanol industry? Appl Microbiol Biotechnol [Internet]. 2013 Feb 28 [cited 2024 Mar 12];97(3):979–91. Available from: https://link.springer.com/article/10.1007/s00253-012-4631-x

67. Raghavendran V, Basso TP, da Silva JB, Basso LC, Gombert AK. A simple scaled down system to mimic the industrial production of first generation fuel ethanol in Brazil. Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology [Internet]. 2017 Jul 1 [cited 2024 Mar 12];110(7):971–83. Available from: https://link.springer.com/article/10.1007/s10482-017-0868-9

68. Hamill PG, Stevenson A, McMullan PE, Williams JP, Lewis ADR, Sudharsan S, et al. Microbial lag phase can be indicative of, or independent from, cellular stress. Scientific Reports 2020 10:1 [Internet]. 2020 Apr 3 [cited 2024 Jan 26];10(1):1–20. Available from: https://www.nature.com/articles/s41598-020-62552-4

69. Rugbjerg P, Sarup-Lytzen K, Nagy M, Sommer MOA. Synthetic addiction extends the productive life time of engineered Escherichia coli populations. Proc Natl Acad Sci U S A [Internet]. 2018 Mar 6 [cited 2024 Jan 26];115(10):2347–52. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.1718622115

70. Rugbjerg P, Sommer MOA. Overcoming genetic heterogeneity in industrial fermentations. Nature Biotechnology 2019 37:8 [Internet]. 2019 Jul 8 [cited 2024 Jan 26];37(8):869–76. Available from: https://www.nature.com/articles/s41587-019-0171-6

71. Rodrigues-Prause A, Sampaio NMV, Gurol TM, Aguirre GM, Sedam HNC, Chapman MJ, et al. A case study of genomic instability in an industrial strain of Saccharomyces cerevisiae. G3: Genes, Genomes, Genetics [Internet]. 2018 Nov 1 [cited 2021 May 25];8(11):3703–13. Available from: https://doi.org/10.1534/g3.118.200446

72. Jiménez-González C, Woodley JM. Bioprocesses: Modeling needs for process evaluation and sustainability assessment. Comput Chem Eng. 2010 Jul 14;34(7):1009–17.

73. Hendratno H. Key Performance Indicators (KPIs). 2021 [cited 2024 Feb 26]; Available from: https://pdc.fide.com/wp-content/uploads/2021/03/Indonesia-report-2020-complete.pdf

74. Taylor C, Marschall L, Kunzelmann M, Richter M, Rudolph F, Vajda J, et al. Integrated process model applications linking bioprocess development to quality by design milestones. Bioengineering [Internet]. 2021 Nov 1 [cited 2024 Feb 26];8(11):156. Available from: https://www.mdpi.com/2306-5354/8/11/156/htm

75. O'Mara P, Farrell A, Bones J, Twomey K. Staying alive! Sensors used for monitoring cell health in bioreactors. Talanta. 2018 Jan 1;176:130–9.

76. Heinzle E. Mass spectrometry for on-line monitoring of biotechnological processes. Adv Biochem Eng Biotechnol [Internet]. 1987 [cited 2024 Jan 30];35:2–45. Available from: https://link.springer.com/chapter/10.1007/BFb0004425

77. Mormino M, Siewers V, Nygård Y. Development of an Haa1-based biosensor for acetic acid sensing in Saccharomyces cerevisiae. FEMS Yeast Res [Internet]. 2021 Sep 1 [cited 2024 Jan 30];21(6). Available from: https://pubmed.ncbi.nlm.nih.gov/34477863/

78. Voisard D, Pugeaud P, Kumar AR, Jenny K, Jayaraman K, Marison IW, et al. Development of a large-scale biocalorimeter to monitor and control bioprocesses. Biotechnol Bioeng [Internet]. 2002 Oct 20 [cited 2024 Jan 27];80(2):125–38. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/bit.10351

79. Cabaneros Lopez P, Feldman H, Mauricio-Iglesias M, Junicke H, Huusom JK, Gernaey K V. Benchmarking real-time monitoring strategies for ethanol production from lignocellulosic biomass. Biomass Bioenergy. 2019 Aug 1;127:105296.

80. Rathore AS, Nikita S, Jesubalan NG. Digitization in bioprocessing: The role of soft sensors in monitoring and control of downstream processing for production of biotherapeutic products. Biosens Bioelectron X. 2022 Dec 1;12:100263.

81. Mukherjee V, Steensels J, Lievens B, Van de Voorde I, Verplaetse A, Aerts G, et al. Phenotypic evaluation of natural and industrial Saccharomyces yeasts for different traits desirable in industrial

bioethanol production. Appl Microbiol Biotechnol [Internet]. 2014 Oct 28 [cited 2024 Jan 26];98(22):9483–98. Available from: https://link.springer.com/article/10.1007/s00253-014-6090-z

82. Balan V. Current Challenges in Commercially Producing Biofuels from Lignocellulosic Biomass. ISRN Biotechnol [Internet]. 2014 May 5 [cited 2024 Jan 26];2014:1–31. Available from: /pmc/articles/PMC4393053/

83. Mohr A, Raman S. Lessons from first generation biofuels and implications for the sustainability appraisal of second generation biofuels. Energy Policy. 2013 Dec 1;63:114–22.

84. Ho DP, Ngo HH, Guo W. A mini review on renewable sources for biofuel. Bioresour Technol. 2014 Oct 1;169:742–9.

85. Zoghlami A, Paës G. Lignocellulosic Biomass: Understanding Recalcitrance and Predicting Hydrolysis. Front Chem. 2019 Dec 18;7:478626.

86. Mankar AR, Pandey A, Modak A, Pant KK. Pretreatment of lignocellulosic biomass: A review on recent advances. Bioresour Technol. 2021 Aug 1;334:125235.

87. Jönsson LJ, Martín C. Pretreatment of lignocellulose: Formation of inhibitory by-products and strategies for minimizing their effects. Bioresour Technol. 2016 Jan 1;199:103–12.

88. Menegazzo F, Ghedini E, Signoretto M. 5-Hydroxymethylfurfural (HMF) Production from Real Biomasses. Molecules : A Journal of Synthetic Chemistry and Natural Product Chemistry [Internet]. 2018 Aug 31 [cited 2024 Jan 26];23(9). Available from: /pmc/articles/PMC6225331/

89. Jönsson LJ, Alriksson B, Nilvebrant NO. Bioconversion of lignocellulose: inhibitors and detoxification. Biotechnol Biofuels [Internet]. 2013 Jan 28 [cited 2018 Sep 28];6(1):16. Available from: http://biotechnologyforbiofuels.biomedcentral.com/articles/10.1186/1754-6834-6-16

90. Sjulander N, Kikas T. Origin, Impact and Control of Lignocellulosic Inhibitors in Bioethanol Production—A Review. Energies 2020, Vol 13, Page 4751 [Internet]. 2020 Sep 11 [cited 2024 Jan 27];13(18):4751. Available from: https://www.mdpi.com/1996-1073/13/18/4751/htm

91. Fache M, Boutevin B, Caillol S. Vanillin Production from Lignin and Its Use as a Renewable Chemical. ACS Sustain Chem Eng [Internet]. 2016 Jan 4 [cited 2024 Jan 27];4(1):35–46. Available from: https://pubs.acs.org/doi/abs/10.1021/acssuschemeng.5b01344

92. Gomes ED, Rodrigues AE. Recovery of vanillin from kraft lignin depolymerization with water as desorption eluent. Sep Purif Technol. 2020 May 15;239:116551.

93. Olsson L, Hahn-Hägerdal B. Fermentation of lignocellulosic hydrolysates for ethanol production. Enzyme Microb Technol. 1996 Apr 1;18(5):312–31.

94. Palmqvist E, Hahn-Hägerdal B. Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. Bioresour Technol. 2000 Aug 1;74(1):25–33.

95. Vanmarcke G, Demeke MM, Foulquié-Moreno MR, Thevelein JM. Identification of the major fermentation inhibitors of recombinant 2G yeasts in diverse lignocellulose hydrolysates. Biotechnol Biofuels [Internet]. 2021 Dec 1 [cited 2023 Sep 19];14(1). Available from: /pmc/articles/PMC8034183/

96. Guo Z, Olsson L. Physiological response of Saccharomyces cerevisiae to weak acids present in lignocellulosic hydrolysate. FEMS Yeast Res [Internet]. 2014 Dec 1 [cited 2022 Mar 21];14(8):1234–48. Available from: https://academic.oup.com/femsyr/article/14/8/1234/2908352

97. Henriques SF, Mira NP, Sá-Correia I. Genome-wide search for candidate genes for yeast robustness improvement against formic acid reveals novel susceptibility (Trk1 and positive regulators) and resistance (Haa1-regulon) determinants. Biotechnol Biofuels [Internet]. 2017 Apr 19 [cited 2022 Mar 21];10(1):1–11. Available from: https://biotechnologyforbiofuels.biomedcentral.com/articles/10.1186/s13068-017-0781-5

98. Oshoma CE, Greetham D, Louis EJ, Smart KA, Phister TG, Powell C, et al. Screening of Non-Saccharomyces cerevisiae Strains for Tolerance to Formic Acid in Bioethanol Fermentation. PLoS One [Internet]. 2015 Aug 18 [cited 2024 Jan 27];10(8). Available from: /pmc/articles/PMC4540574/

99. Guo ZP, Khoomrung S, Nielsen J, Olsson L. Changes in lipid metabolism convey acid tolerance in Saccharomyces cerevisiae 06 Biological Sciences 0601 Biochemistry and Cell Biology. Biotechnol Biofuels [Internet]. 2018 Oct 29 [cited 2024 Jan 27];11(1):1–15. Available from: https://biotechnologyforbiofuels.biomedcentral.com/articles/10.1186/s13068-018-1295-5

100. Beckner M, Ivey ML, Phister TG. Microbial contamination of fuel ethanol fermentations. Lett Appl Microbiol [Internet]. 2011 [cited 2024 Jan 27];53(4):387–94. Available from: https://pubmed.ncbi.nlm.nih.gov/21770989/

101. Peetermans A, Foulquié-Moreno MR, Thevelein JM. Mechanisms underlying lactic acid tolerance and its influence on lactic acid production in Saccharomyces cerevisiae. Microbial Cell [Internet]. 2021 Jun 1 [cited 2022 Feb 22];8(6):111. Available from: /pmc/articles/PMC8144909/

102. Senne de Oliveira Lino F, Bajic D, Vila JCC, Sánchez A, Sommer MOA. Complex yeast–bacteria interactions affect the yield of industrial ethanol fermentation. Nature Communications 2021 12:1 [Internet]. 2021 Mar 8 [cited 2021 Mar 9];12(1):1–12. Available from: https://www.nature.com/articles/s41467-021-21844-7

103. Taherzadeh MJ, Gustafsson L, Niklasson C, Lidén G. Physiological effects of 5-hydroxymethylfurfural on Saccharomyces cerevisiae. Applied Microbiology and Biotechnology 2000 53:6 [Internet]. 2000 [cited 2022 Feb 22];53(6):701–8. Available from: https://link.springer.com/article/10.1007/s002530000328

104. Wang X, Liang Z, Hou J, Bao X, Shen Y. Identification and functional evaluation of the reductases and dehydrogenases from Saccharomyces cerevisiae involved in vanillin resistance. BMC Biotechnol [Internet]. 2016 Apr 1 [cited 2024 Feb 27];16(1):1–9. Available from: https://bmcbiotechnol.biomedcentral.com/articles/10.1186/s12896-016-0264-y

105. Iwaki A, Ohnuki S, Suga Y, Izawa S, Ohya Y. Vanillin Inhibits Translation and Induces Messenger Ribonucleoprotein (mRNP) Granule Formation in Saccharomyces cerevisiae: Application and Validation of High-Content, Image-Based Profiling. PLoS One [Internet]. 2013 Apr 24 [cited 2024 Jan 27];8(4):e61748. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0061748

106. Lindahl L, Genheden S, Eriksson LA, Olsson L, Bettiga M. Sphingolipids contribute to acetic acid resistance in *Zygosaccharomyces bailii*. Biotechnol Bioeng [Internet]. 2016 Apr 1 [cited 2019 Mar 22];113(4):744–53. Available from: http://doi.wiley.com/10.1002/bit.25845

107. Lindahl L, Santos AXS, Olsson H, Olsson L, Bettiga M. Membrane engineering of S. cerevisiae targeting sphingolipid metabolism. Sci Rep [Internet]. 2017 Dec 1 [cited 2018 Oct 1];7(1):41868. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28145511

108. Nadal-Rey G, McClure DD, Kavanagh JM, Cornelissen S, Fletcher DF, Gernaey K V. Understanding gradients in industrial bioreactors. Biotechnol Adv. 2021 Jan 1;46:107660.

109. Mitra S, Murthy GS. Bioreactor control systems in the biopharmaceutical industry: a critical perspective. Systems Microbiology and Biomanufacturing [Internet]. 2022 Jan 1 [cited 2024 Jan 30];2(1):91. Available from: /pmc/articles/PMC8340809/

110. Deparis Q, Claes A, Foulquié-Moreno MR, Thevelein JM. Engineering tolerance to industrially relevant stress factors in yeast cell factories [Internet]. Oxford Academic; Jun 1, 2017 p. 36. Available from: https://academic.oup.com/femsyr/article/17/4/fox036/3861662

111. Woo JM, Yang KM, Kim SU, Blank LM, Park JB. High temperature stimulates acetic acid accumulation and enhances the growth inhibition and ethanol production by Saccharomyces cerevisiae under fermenting conditions. Appl Microbiol Biotechnol [Internet]. 2014 Apr 5 [cited 2024 Mar 12];98(13):6085–94. Available from: https://link.springer.com/article/10.1007/s00253-014-5691-x

112. Westman JO, Wang R, Novy V, Franzén CJ. Sustaining fermentation in high-gravity ethanol production by feeding yeast to a temperature-profiled multifeed simultaneous saccharification and co-fermentation of wheat straw. Biotechnol Biofuels [Internet]. 2017 Sep 12 [cited 2024 Mar 6];10(1):1–16. Available from: https://biotechnologyforbiofuels.biomedcentral.com/articles/10.1186/s13068-017-0893-y

113. Costanzo M, Hou J, Messier V, Nelson J, Rahman M, VanderSluis B, et al. Environmental robustness of the global yeast genetic interaction network. Science (1979) [Internet]. 2021 May 7 [cited 2022 Oct 3];372(6542). Available from: https://www.science.org/doi/10.1126/science.abf8424

114. Persson K, Stenberg S, Tamás MJ, Warringer J. Adaptation of the yeast gene knockout collection is near-perfectly predicted by fitness and diminishing return epistasis. G3 Genes|Genomes|Genetics [Internet]. 2022 Nov 4 [cited 2024 Jan 23];12(11). Available from: https://dx.doi.org/10.1093/g3journal/jkac240

115. Carlson JM, Doyle J. Complexity and robustness. Proc Natl Acad Sci U S A [Internet]. 2002 Feb 19 [cited 2024 Jan 31];99(SUPPL. 1):2538–45. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.012582499

116. Xu H, Mannor S. Robustness and generalization. Mach Learn [Internet]. 2012 Mar 15 [cited 2024 Jan 31];86(3):391–423. Available from: https://link.springer.com/article/10.1007/s10994-011-5268-1

117. Grimbs S, Selbig J, Bulik S, Holzhütter HG, Steuer R. The stability and robustness of metabolic states: Identifying stabilizing sites in metabolic networks. Mol Syst Biol [Internet]. 2007 Nov 13 [cited 2024 Jan 31];3. Available from: https://www.embopress.org/doi/10.1038/msb4100186

118. Holme P. Metabolic Robustness and Network Modularity: A Model Study. PLoS One [Internet]. 2011 [cited 2024 Jan 31];6(2). Available from: /pmc/articles/PMC3032788/

119. Fares MA. The origins of mutational robustness. Trends Genet [Internet]. 2015 Jul 1 [cited 2024 Jan 31];31(7):373–81. Available from: https://pubmed.ncbi.nlm.nih.gov/26013677/

120. Wilke CO. Transcriptional Robustness Complements Nonsense-Mediated Decay in Humans. PLoS Genet [Internet]. 2011 Oct [cited 2024 Jan 31];7(10):e1002296. Available from: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1002296

121. Fisher AC, Kamga MH, Agarabi C, Brorson K, Lee SL, Yoon S. The Current Scientific and Regulatory Landscape in Advancing Integrated Continuous Biopharmaceutical Manufacturing. Trends Biotechnol. 2019 Mar 1;37(3):253–67.

122. Tian T, Olson S, Whitacre JM, Harding A. The origins of cancer robustness and evolvability. Integr Biol (Camb) [Internet]. 2011 Jan [cited 2023 Dec 13];3(1):17–30. Available from: https://pubmed.ncbi.nlm.nih.gov/20944865/

123. Kitano H. Cancer robustness: Tumour tactics. Nature 2003 426:6963 [Internet]. 2003 Nov 13 [cited 2023 Feb 1];426(6963):125–125. Available from: https://www.nature.com/articles/426125a

124. Hernando-Amado S, Laborda P, Martínez JL. Tackling antibiotic resistance by inducing transient and robust collateral sensitivity. Nature Communications 2023 14:1 [Internet]. 2023 Mar 30 [cited 2024 Feb 1];14(1):1–12. Available from: https://www.nature.com/articles/s41467-023-37357-4

125. Pang Z, Raudonis R, Glick BR, Lin TJ, Cheng Z. Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies. Biotechnol Adv. 2019 Jan 1;37(1):177–92.

126. Torello Pianale L, Caputo F, Olsson L. Four ways of implementing robustness quantification in strain characterisation. Biotechnology for Biofuels and Bioproducts [Internet]. 2023 Dec 1 [cited 2024 Feb 1];16(1):1–20. Available from: https://link.springer.com/articles/10.1186/s13068-023-02445-6

127. Blöbaum L, Torello Pianale L, Olsson L, Grünberger A. Quantifying microbial robustness in dynamic environments using microfluidic single-cell cultivation. Microb Cell Fact [Internet]. 2024 Dec 1 [cited 2024 Mar 4];23(1):44. Available from: https://microbialcellfactories.biomedcentral.com/articles/10.1186/s12934-024-02318-z

128. Trivellin C, Rugbjerg P, Olsson L. Performance and robustness analysis reveals phenotypic trade-offs in yeast. Life Sci Alliance [Internet]. 2024 Jan 1;7(1):e202302215. Available from: https://www.life-science-alliance.org/content/7/1/e202302215.abstract

129. Yang L, Srinivasan S, Mahadevan R, Cluett WR. Characterizing metabolic pathway diversification in the context of perturbation size. Metab Eng [Internet]. 2015 Mar 1 [cited 2022 Mar 22];28:114–22. Available from: https://pubmed.ncbi.nlm.nih.gov/25542850/

130. Bianconi F, Antonini C, Tomassoni L, Valigi P. CRA toolbox: Software package for conditional robustness analysis of cancer systems biology models in MATLAB. BMC Bioinformatics [Internet]. 2019 Jul 9 [cited 2021 Jan 28];20(1):385. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-019-2933-z

131. Bianconi F, Baldelli E, Luovini V, Petricoin EF, Crinò L, Valigi P. Conditional robustness analysis for fragility discovery and target identification in biochemical networks and in cancer systems biology. BMC Syst Biol [Internet]. 2015 Dec 19 [cited 2021 Jan 28];9(1):70. Available from: http://bmcsystbiol.biomedcentral.com/articles/10.1186/s12918-015-0216-5

132. Rajdl K, Lansky P, Kostal L. Fano Factor: A Potentially Useful Information. Front Comput Neurosci. 2020 Nov 20;0:100.

133. Zheng L, Chen M, Nie Q. External noise control in inherently stochastic biological systems. J Math Phys [Internet]. 2012 Nov 27 [cited 2021 Oct 14];53(11):115616. Available from: /pmc/articles/PMC3505197/

134. Balderas-Hernández VE, Correia K, Mahadevan R. Inactivation of the transcription factor mig1 (YGL035C) in Saccharomyces cerevisiae improves tolerance towards monocarboxylic weak acids: acetic, formic and levulinic acid. J Ind Microbiol Biotechnol [Internet]. 2018 Aug 1 [cited 2022 Jan 14];45(8):735–51. Available from: https://academic.oup.com/jimb/article/45/8/735/5996672

135. Baym M, Lieberman TD, Kelsic ED, Chait R, Gross R, Yelin I, et al. Spatiotemporal microbial evolution on antibiotic landscapes. Science [Internet]. 2016 Sep 9 [cited 2024 Feb 1];353(6304):1147. Available from: /pmc/articles/PMC5534434/

136. Schoborg JA, Clark LG, Choudhury A, Hodgman CE, Jewett MC. Yeast knockout library allows for efficient testing of genomic mutations for cell-free protein synthesis. Synth Syst Biotechnol. 2016 Mar 1;1(1):2–6.

137. Almeida LD, Silva ASF, Mota DC, Vasconcelos AA, Camargo AP, Pires GS, et al. Yeast Double Transporter Gene Deletion Library for Identification of Xenobiotic Carriers in Low or High Throughput. mBio [Internet]. 2021 Dec 1 [cited 2023 Aug 21];12(6). Available from: https://journals.asm.org/doi/10.1128/mbio.03221-21

138. Martínez VS, Nielsen LK. NExT: Integration of thermodynamic constraints and metabolomics data into a metabolic network. Methods in Molecular Biology [Internet]. 2014 [cited 2024 Feb 7];1191:65–78. Available from: https://link.springer.com/protocol/10.1007/978-1-4939-1170-7_4

139. Saltz JB, Hessel FC, Kelly MW. Trait Correlations in the Genomics Era. Trends Ecol Evol. 2017 Apr 1;32(4):279–90.

140. Rugbjerg P, Olsson L. The future of self-selecting and stable fermentations. 2020 Nov [cited 2021 Nov 5];47(11):993–1004. Available from: https://link.springer.com/article/10.1007/s10295-020-02325-0

141. Lovero KG, Treseder KK. Trade-Offs Between Growth Rate and Other Fungal Traits. Frontiers in Forests and Global Change. 2021 Nov 30;4:197.

142. Basan M, Honda T, Christodoulou D, Hörl M, Chang YF, Leoncini E, et al. A universal trade-off between growth and lag in fluctuating environments. Nature 2020 584:7821 [Internet]. 2020 Jul 15 [cited 2022 Oct 2];584(7821):470–4. Available from: /pmc/articles/PMC7442741/

143. Pfeiffer T, Schuster S, Bonhoeffer S. Cooperation and competition in the evolution of ATP-producing pathways. Science (1979) [Internet]. 2001 Apr 20 [cited 2024 Feb 6];292(5516):504–7. Available from: https://www.science.org/doi/10.1126/science.1058079

144. Wei X, Zhang J. Environment-dependent pleiotropic effects of mutations on the maximum growth rate r and carrying capacity K of population growth. PLoS Biol [Internet]. 2019 [cited 2024 Feb 5];17(1). Available from: /pmc/articles/PMC6364931/

145. Patanè A, Santoro A, Costanza J, Carapezza G, Nicosia G. Pareto Optimal Design for Synthetic Biology. IEEE Trans Biomed Circuits Syst. 2015 Aug 1;9(4):555–71.

146. Shoval O, Sheftel H, Shinar G, Hart Y, Ramote O, Mayo A, et al. Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space. Science (1979) [Internet]. 2012 Jun 1 [cited 2022 Oct 11];336(6085):1157–60. Available from: https://www.science.org/doi/10.1126/science.1217405

147. Sheftel H, Shoval O, Mayo A, Alon U. The geometry of the Pareto front in biological phenotype space. Ecol Evol [Internet]. 2013 Jun 1 [cited 2022 Aug 18];3(6). Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.528

148. Legon L, Corre C, Bates DG, Mannan AA. gcFront: a tool for determining a Pareto front of growth-coupled cell factory designs. Bioinformatics [Internet]. 2022 Jul 11 [cited 2022 Sep 8];38(14):3657–9. Available from: https://academic.oup.com/bioinformatics/article/38/14/3657/6596618

149. Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U. Multidimensional optimality of microbial metabolism. Science (1979) [Internet]. 2012 May 4 [cited 2023 Mar 1];336(6081):601–4. Available from: https://www.science.org/doi/10.1126/science.1216882

150. Li Y, Petrov DA, Sherlock G. Single nucleotide mapping of trait space reveals Pareto fronts that constrain adaptation. Nature Ecology & Evolution 2019 3:11 [Internet]. 2019 Oct 14 [cited 2024 Feb 15];3(11):1539–51. Available from: https://www.nature.com/articles/s41559-019-0993-0

151. Krakauer DC, Plotkin JB. Redundancy, antiredundancy, and the robustness of genomes. Proc Natl Acad Sci U S A [Internet]. 2002 Feb 5 [cited 2024 Feb 12];99(3):1405–9. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.032668599

152. MacNeil LT, Walhout AJM. Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. Genome Res [Internet]. 2011 [cited 2024 Feb 12];21(5):645. Available from: /pmc/articles/PMC3083081/

153. Alcalá-Corona SA, Sandoval-Motta S, Espinal-Enríquez J, Hernández-Lemus E. Modularity in Biological Networks. Front Genet [Internet]. 2021 Sep 14 [cited 2024 Feb 12];12:701331. Available from: www.frontiersin.org

154. McBride CD, Grunberg TW, Del Vecchio D. Design of genetic circuits that are robust to resource competition. Curr Opin Syst Biol. 2021 Dec 1;28:100357.

155. Ingolia NT, Murray AW. Positive feedback loops as a flexible biological module. Curr Biol [Internet]. 2007 Apr 4 [cited 2024 Feb 12];17(8):668. Available from: /pmc/articles/PMC1914375/

156. Schopf FH, Biebl MM, Buchner J. The HSP90 chaperone machinery. Nature Reviews Molecular Cell Biology 2017 18:6 [Internet]. 2017 Apr 21 [cited 2024 Feb 12];18(6):345–60. Available from: https://www.nature.com/articles/nrm.2017.20

157. Girstmair H, Tippel F, Lopez A, Tych K, Stein F, Haberkant P, et al. The Hsp90 isoforms from S. cerevisiae differ in structure, function and client range. Nature Communications 2019 10:1 [Internet]. 2019 Aug 9 [cited 2024 Feb 12];10(1):1–15. Available from: https://www.nature.com/articles/s41467-019-11518-w

158. Sangster TA, Lindquist S, Queitsch C. Under cover: Causes, effects and implications of Hsp90-mediated genetic capacitance. BioEssays. 2004 Apr;26(4):348–62.

159. Levy SF, Siegal ML. Network Hubs Buffer Environmental Variation in Saccharomyces cerevisiae. PLoS Biol [Internet]. 2008 Nov [cited 2024 Feb 12];6(11):e264. Available from: https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0060264

160. Mináriková L, Kuthan M, Ricicová M, Forstová J, Palková Z. Differentiated Gene Expression in Cells within Yeast Colonies. Exp Cell Res. 2001 Dec 10;271(2):296–304.

161. Manukyan A, Zhang J, Thippeswamy U, Yang J, Zavala N, Mudannayake MP, et al. Ccr4 Alters Cell Size in Yeast by Modulating the Timing of CLN1 and CLN2 Expression. Genetics [Internet]. 2008 May 1 [cited 2024 Feb 12];179(1):345–57. Available from: https://dx.doi.org/10.1534/genetics.108.086744

162. Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, et al. Regulated cell-to-cell variation in a cell-fate decision system. Nature 2005 437:7059 [Internet]. 2005 Sep 18 [cited 2024 Feb 12];437(7059):699–706. Available from: https://www.nature.com/articles/nature03998

163. Cooper TF, Morby AP, Gunn A, Schneider D. Effect of random and hub gene disruptions on environmental and mutational robustness in Escherichia coli. BMC Genomics [Internet]. 2006 Sep 18 [cited 2024 Feb 12];7:237. Available from: /pmc/articles/PMC1590030/

164. Milton CC, Huynh B, Batterham P, Rutherford SL, Hoffmann AA. Quantitative trait symmetry independent of Hsp90 buffering: Distinct modes of genetic canalization and developmental stability. Proceedings of the National Academy of Sciences [Internet]. 2003 Nov 11 [cited 2024 Feb 12];100(23):13396–401. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.1835613100

165. Taymaz-Nikerel H, Cankorur-Cetinkaya A, Kirdar B. Genome-wide transcriptional response of Saccharomyces cerevisiae to stress-induced perturbations. Front Bioeng Biotechnol. 2016 Feb 18;4(FEB):167523.

166. Ohya Y, Sese J, Yukawa M, Sano F, Nakatani Y, Saito TL, et al. High-dimensional and large-scale phenotyping of yeast mutants. Proc Natl Acad Sci U S A [Internet]. 2005 Dec 27 [cited 2024 Feb 12];102(52):19015–20. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.0509436102

167. Kuras L, Barbey R, Thomas D. Assembly of a bZIP-bHLH transcription activation complex: formation of the yeast Cbf1-Met4-Met28 complex is regulated through Met28 stimulation of Cbf1 DNA binding. EMBO J [Internet]. 1997 May 5 [cited 2023 Aug 24];16(9):2441. Available from: /pmc/articles/PMC1169844/?report=abstract

168. Kuras L, Cherest H, Surdin-Kerjan Y, Thomas D. A heteromeric complex containing the centromere binding factor 1 and two basic leucine zipper factors, Met4 and Met28, mediates the transcription activation of yeast sulfur metabolism. EMBO J [Internet]. 1996 May 5 [cited 2023 Oct 2];15(10):2519. Available from: /pmc/articles/PMC450184/?report=abstract

169. Raghavendran V, Marx C, Olsson L, Bettiga M. The protective role of intracellular glutathione in Saccharomyces cerevisiae during lignocellulosic ethanol production. AMB Express [Internet]. 2020 Dec 1 [cited 2024 Feb 12];10(1):1–8. Available from: https://amb-express.springeropen.com/articles/10.1186/s13568-020-01148-7

170. Grant CM, MacIver FH, Dawes IW. Glutathione is an essential metabolite required for resistance to oxidative stress in the yeast Saccharomyces cerevisiae. Curr Genet [Internet]. 1996 [cited 2024 Feb 12];29(6):511–5. Available from: https://pubmed.ncbi.nlm.nih.gov/8662189/

171. Turco G, Chang C, Wang RY, Kim G, Stoops EH, Richardson B, et al. Global analysis of the yeast knockout phenome. Sci Adv [Internet]. 2023 May 26 [cited 2024 Feb 12];9(21). Available from: https://www.science.org/doi/10.1126/sciadv.adg5702

172. Zackrisson M, Hallin J, Ottosson LG, Dahl P, Fernandez-Parada E, Ländström E, et al. Scan-o-matic: High-resolution microbial phenomics at a massive scale. G3: Genes, Genomes, Genetics [Internet]. 2016 [cited 2023 Nov 8];6(9):3003–14. Available from: /pmc/articles/PMC5015956/

173. Alic N, Higgins VJ, Dawes IW. Identification of a Saccharomyces cerevisiae Gene that Is Required for G1 Arrest in Response to the Lipid Oxidation Product Linoleic Acid Hydroperoxide*. Mol Biol Cell [Internet]. 2001 [cited 2024 Feb 18];12(6):1801. Available from: /pmc/articles/PMC37342/

174. Tabb-Massey A, Caffrey JM, Logsden P, Taylor S, Trent JO, Ellis SR. Ribosomal proteins Rps0 and Rps21 of Saccharomyces cerevisiae have overlapping functions in the maturation of the 3′ end of 18S rRNA. Nucleic Acids Res [Internet]. 2003 Dec 1 [cited 2024 Feb 18];31(23):6798–805. Available from: https://dx.doi.org/10.1093/nar/gkg899

175. Gierest H, Thao NN, Surdin-Kerjan Y. Transcriptional regulation of the MET3 gene of Saccharomyces cerevisiae. Gene. 1985 Jan 1;34(2–3):269–81.

176. Bush GL, Meyer DI. The refolding activity of the yeast heat shock proteins Ssa1 and Ssa2 defines their role in protein translocation. J Cell Biol [Internet]. 1996 Dec 12 [cited 2024 Feb 19];135(5):1229. Available from: /pmc/articles/PMC2121094/?report=abstract

177. Kater L, Thoms M, Barrio-Garcia C, Cheng J, Ismail S, Ahmed YL, et al. Visualizing the Assembly Pathway of Nucleolar Pre-60S Ribosomes. Cell. 2017 Dec 14;171(7):1599-1610.e14.

178. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced

genome of interest. Nucleic Acids Res [Internet]. 2023 Jan 6 [cited 2024 Feb 21];51(D1):D638–46. Available from: https://dx.doi.org/10.1093/nar/gkac1000

179. Mavrommati M, Daskalaki A, Papanikolaou S, Aggelis G. Adaptive laboratory evolution principles and applications in industrial biotechnology. Biotechnol Adv. 2022 Jan 1;54:107795.

180. Loewe L, Hill WG. The population genetics of mutations: good, bad and indifferent. Philosophical Transactions of the Royal Society B: Biological Sciences [Internet]. 2010 Apr 4 [cited 2024 Feb 21];365(1544):1153. Available from: /pmc/articles/PMC2871823/

181. Desai MM, Fisher DS. Beneficial Mutation–Selection Balance and the Effect of Linkage on Positive Selection. Genetics [Internet]. 2007 Jul 1 [cited 2024 Feb 21];176(3):1759–98. Available from: https://dx.doi.org/10.1534/genetics.106.067678

182. Zambrano MM, Kolter R. GASPing for life in stationary phase. Cell [Internet]. 1996 Jul 26 [cited 2024 Feb 21];86(2):181–4. Available from: http://www.cell.com/article/S0092867400800896/fulltext

183. Wright J, Bellissimi E, de Hulster E, Wagner A, Pronk JT, Van Maris AJA. Batch and continuous culture-based selection strategies for acetic acid tolerance in xylose-fermenting Saccharomyces cerevisiae. FEMS Yeast Res [Internet]. 2011 May 1 [cited 2024 Feb 21];11(3):299–306. Available from: https://dx.doi.org/10.1111/j.1567-1364.2011.00719.x

184. Weikert C, Sauer U, Bailey JE. Use of a glycerol-limited, long-term chemostat for isolation of Escherichia coli mutants with improved physiological properties. Microbiology (N Y) [Internet]. 1997 May 1 [cited 2024 Feb 21];143(5):1567–74. Available from: https://www.microbiologyresearch.org/content/journal/micro/10.1099/00221287-143-5-1567

185. Bono LM, Smith LB, Pfennig DW, Burch CL. The emergence of performance trade-offs during local adaptation: insights from experimental evolution. Mol Ecol [Internet]. 2017 Apr 1 [cited 2024 Feb 21];26(7):1720–33. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/mec.13979

186. Bono LM, Draghi JA, Turner PE. Evolvability Costs of Niche Expansion. Trends in Genetics [Internet]. 2020 Jan 1 [cited 2024 Feb 21];36(1):14–23. Available from: http://www.cell.com/article/S0168952519301921/fulltext

187. Dhar R, Sägesser R, Weikert C, Wagner A. Yeast adapts to a changing stressful environment by evolving cross-protection and anticipatory gene regulation. Mol Biol Evol [Internet]. 2013 Mar 1 [cited 2023 Jan 10];30(3):573–88. Available from: https://dx.doi.org/10.1093/molbev/mss253

188. Fasanello VJ, Liu P, Fay JC, Botero CA. Fluctuating selection facilitates the discovery of broadly effective but difficult to reach adaptive outcomes in yeast. Evol Lett [Internet]. 2023 Nov 10 [cited 2024 Feb 21];1–10. Available from: https://dx.doi.org/10.1093/evlett/qrad055

189. Wiser MJ, Lenski RE. A Comparison of Methods to Measure Fitness in Escherichia coli. PLoS One [Internet]. 2015 May 11 [cited 2024 Feb 21];10(5). Available from: /pmc/articles/PMC4427439/

190. Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ. Empirical fitness landscapes reveal accessible evolutionary paths. Nature 2006 445:7126 [Internet]. 2007 Jan 25 [cited 2024 Feb 22];445(7126):383–6. Available from: https://www.nature.com/articles/nature05451

191. Payne JL, Wagner A. The causes of evolvability and their evolution. Nature Reviews Genetics 2018 20:1 [Internet]. 2018 Nov 1 [cited 2024 Feb 22];20(1):24–38. Available from: https://www.nature.com/articles/s41576-018-0069-z

192. Couce A, Tenaillon OA. The rule of declining adaptability in microbial evolution experiments. Front Genet [Internet]. 2015 Mar 11 [cited 2024 Feb 22];6(MAR):128797. Available from: www.frontiersin.org

193. Nelson P, Masel J. Evolutionary Capacitance Emerges Spontaneously during Adaptation to Environmental Changes. Cell Rep [Internet]. 2018 Oct 2 [cited 2024 Feb 22];25(1):249–58. Available from: https://pubmed.ncbi.nlm.nih.gov/30282033/

194. Newnam GP, Wegrzyn RD, Lindquist SL, Chernoff YO. Antagonistic Interactions between Yeast Chaperones Hsp104 and Hsp70 in Prion Curing. Mol Cell Biol [Internet]. 1999 Feb 1 [cited 2024 Feb 22];19(2):1325. Available from: /pmc/articles/PMC116061/

195. Wickner RB, Shewmaker FP, Bateman DA, Edskes HK, Gorkovskiy A, Dayani Y, et al. Yeast prions: structure, biology, and prion-handling systems. Microbiol Mol Biol Rev [Internet]. 2015 Mar [cited 2024 Mar 9];79(1):1–17. Available from: https://pubmed.ncbi.nlm.nih.gov/25631286/

196. Singh K, Pil JK, Park HO. The Rho5 GTPase is necessary for oxidant-induced cell death in budding yeast. Proc Natl Acad Sci U S A [Internet]. 2008 Feb 5 [cited 2024 Feb 22];105(5):1522–7. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.0707359105

197. Ryan DP, Owen-Hughes T. Snf2-family proteins: chromatin remodellers for any occasion. Curr Opin Chem Biol [Internet]. 2011 Oct [cited 2024 Feb 22];15(5):649. Available from: /pmc/articles/PMC4162295/

198. Grimminger-Marquardt V, Lashuel HA. Structure and function of the molecular chaperone Hsp104 from yeast. Biopolymers [Internet]. 2010 [cited 2024 Mar 9];93(3):252–76. Available from: https://pubmed.ncbi.nlm.nih.gov/19768774/

199. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. Nature Reviews Methods Primers 2021 1:1 [Internet]. 2021 Aug 26 [cited 2024 Feb 27];1(1):1–21. Available from: https://www.nature.com/articles/s43586-021-00056-9

200. Powder KE. Quantitative Trait Loci (QTL) Mapping. Methods Mol Biol [Internet]. 2020 [cited 2024 Feb 27];2082:211–29. Available from: https://pubmed.ncbi.nlm.nih.gov/31849018/

201. Vermeersch L, Perez-Samper G, Cerulus B, Jariani A, Gallone B, Voordeckers K, et al. On the duration of the microbial lag phase. Curr Genet [Internet]. 2019 Jun 1 [cited 2024 Feb 23];65(3):721. Available from: /pmc/articles/PMC6510831/

202. Vermeersch L, Cool L, Gorkovskiy A, Voordeckers K, Wenseleers T, Verstrepen KJ. Do microbes have a memory? History-dependent behavior in the adaptation to variable environments. Front Microbiol [Internet]. 2022 Oct 10 [cited 2024 Mar 9];13. Available from: /pmc/articles/PMC9589428/

203. Bisgaard J, Muldbak M, Cornelissen S, Tajsoleiman T, Huusom JK, Rasmussen T, et al. Flow-following sensor devices: A tool for bridging data and model predictions in large-scale fermentations. Comput Struct Biotechnol J. 2020 Jan 1;18:2908–19.

204. Delvigne F, Takors R, Mudde R, van Gulik W, Noorman H. Bioprocess scale-up/down as integrative enabling technology: from fluid mechanics to systems biology and beyond. Microb Biotechnol [Internet]. 2017 Sep 1 [cited 2022 Mar 22];10(5):1267–74. Available from: https://pubmed.ncbi.nlm.nih.gov/28805306/

205. Pinheiro S, Pandey S, Pelet S. Cellular heterogeneity: yeast-side story. Fungal Biol Rev. 2022 Mar 1;39:34–45.